

AD-A047 056

CORNELL UNIV ITHACA N Y SCHOOL OF OPERATIONS RESEARC--ETC F/6 12/1
DESIGN OF EXPERIMENTS FOR SELECTION FROM ORDERED FAMILIES OF DI--ETC(U)
AUG 77 J H HOOPER, T J SANTNER

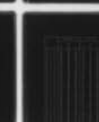
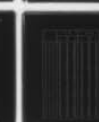
DAA629-73-C-0008

UNCLASSIFIED

TR-350

NL

1 9f 1
AD
A047 056



END

DATE

FILMED

1 - 78

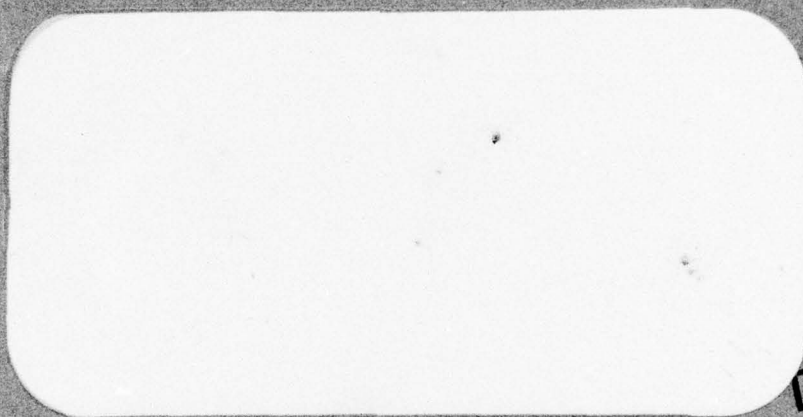
DDC

AD A 0 47056

SCHOOL
OF
OPERATIONS RESEARCH
AND
INDUSTRIAL ENGINEERING



(12)
B.S.



DDC
RECEIVED
NOV 23 1977
F



COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853

AD No. _____
DDC FILE COPY

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

(12)

SCHOOL OF OPERATIONS RESEARCH
AND INDUSTRIAL ENGINEERING
COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NEW YORK

(14) TR-350

(9) TECHNICAL REPORT NO. 350

(11) Aug 1977

(12) 55p.

(6) Design of Experiments for Selection from
Ordered Families of Distributions.

by

(10) Jeffrey H. Hooper

Thomas J. Santner



(21) Also available under contract nos.
NSF-ENG-75-10487 and N00014-75-C-0586.

(15)
Prepared under Contracts National Science Foundation ENG75-10487,
U.S. Army Research Office DAAG29-73-C-0008, DAAG29-77-C-0003 Office
of Naval Research N00014-75-C-0586.

Approved for Public Release; Distribution Unlimited.

409869

Inc

1. Introduction

Given a space S of cumulative distribution functions (cdf's) and a weak (reflective and transitive) order relation, $<$, on S , a subclass of S is called an ordered family of distributions (with respect to $G \in S$) if it has the form $\{F \in S \mid F < G\}$. Two familiar examples of weak orderings are convex ordering and star ordering which yield the increasing failure rate (IFR) and increasing failure rate on the average (IFRA) families when $G(x) = 1 - e^{-x}$.

Weak orderings and the ordered family model assumption have proved to be of considerable importance in the literature: (a) for providing characterizations of skewness and kurtosis (Van Zwet (1964)) (b) in power studies of certain rank tests (Doksum (1969)) as models for the lifetimes of coherent systems (Birnbaum, Esuery, and Marshall (1966)) and (d) as models for the lifetimes of systems subject to random shocks (Esuery, Marshall and Proschan (1973) and Barlow and Proschan (1975)).

→ This paper studies the problem of designing a single stage experiment for the selection of the distribution(s) with the largest α -quantile(s) where $\alpha \in (0,1)$ is given and when the model assumptions are specified by an ordered family of distributions. More precisely suppose observations are to be taken from k populations labeled Π_1, \dots, Π_k . For each i in $\{1, \dots, k\}$ let (a) F_i denote the cdf of population Π_i , (b) $\{X_{ij} : 1 \leq j \leq n\}$ be a random sample from Π_i , (c) $x_\alpha(F_i)$ be the α -quantile of F_i , (d) $F_{[i]}$ be the cdf with the i th smallest α -quantile, and (e) $\Pi_{(i)}$ denote the (unknown) population with cdf $F_{[i]}$. Then $x_\alpha(F_{[i]}) \leq \dots \leq x_\alpha(F_{[k]})$ are the ordered α -quantiles. Assume samples from different Π_i 's are independent, there is no prior knowledge of the

pairing of Π_i and $\Pi_{(i)}$, the F_i 's come from a specified ordered family of distributions, and the experimenter is interested in selection of $\Pi_{(k-t+1)}, \dots, \Pi_{(k)}$, the t populations with largest α -quantiles. We adopt the preferred population formulation in which Π_i is called a preferred population if $x_\alpha(F_i)$ is "close" (in a sense made precise in Section 3) to $x_\alpha(F_{[k-t+1]})$ and consider two classes of selection problems:

1. selection of a subset of prespecified size $s(r \leq t, r \leq s \leq k-t+r)$ so as to include at least r preferred populations and
2. selection of a random size subset containing at most $s(1 \leq s \leq k)$ populations so as to include at least one ($r=1$) preferred population.

The problem is to design the experiment so that (1) or (2) is guaranteed with at least prespecified probability P^* no matter what the true vector of cdf's. This approach combines aspects of Gupta's subset selection formulation (1956, 65) and Bechhofer's indifference zone formulation (1954) and will be shown to be a strengthening of the latter.

Previous work on selection procedures for nonparametric problems can be categorized according to (1) the parameter of interest for selection, (2) the formulation of the design requirement and (3) the model assumptions. One group of papers considers problems of selection for location parameters or more complicated linear models usually under the assumption that the error distributions are continuous and symmetric. They include: Lehmann (1963), Bartlett and Govindarajulu (1968), Randles (1970), Bhapkar and Gore (1971), Puri and Sen (1972), Ghosh (1973) and Gupta and Huang (1974). Related papers include Patel (1976) who considers the problem of selecting the distribution with the largest scale parameter when the common form of the distribution is assumed to be IFR and Gupta and McDonald

(1969) who consider selection in terms of an arbitrary parameter under which the distributions form a stochastically increasing family of distributions. A second group of papers are those which consider selection in terms of α -quantiles. Rizvi and Sobel (1967), Rizvi, Sobel and Woodworth (1968) and Barlow and Gupta (1969) study subset selection formulations of the problem while Sobel (1967) and Desu and Sobel (1971) use an indifference zone formulation. For a complete review of the literature see Lee and Dudewicz (1974).

In contrast, the present work does not assume the F_i 's are stochastically ordered by a location, scale or other parameter as do the papers in the first group nor does it assume they are arbitrary continuous cdf's as do many of the papers in the second group. Rather the ordered family assumption falls between these two extremes.

Section 2 studies weak orderings and the corresponding ordered families. Sections 3 and 4 formulate a fixed size subset selection problem and a random size subset selection problem respectively and give their solution for two classes of distributions. Finally Section 5 describes some properties of the two classes of procedures considered in Sections 3 and 4.

ACCESSION for	
NTIS	W. Po Section <input checked="" type="checkbox"/>
DDC	Bull Section <input type="checkbox"/>
UNANNOUNCED	<input type="checkbox"/>
JUSTIFICATION	
BY	
DISTRIBUTION/AVAILABILITY CODES	
DI	SPECIAL
A	

2. Weak Orderings and Ordered Families of Distributions

Given a finite or infinite interval, I , of the real line, R , let F_I denote the set of distributions F which satisfy:

- (1) F is absolutely continuous with respect to Lebesgue measure,
- (2) $S(F) \equiv \{x \in R \mid F(x+\epsilon) - F(x-\epsilon) > 0 \forall \epsilon > 0\}$, the support of F , is a convex set,
- (3) $S(F) \subset I$.

In particular when $I = R$ and $[0, \infty)$, F_I will be denoted as F and F_0 respectively. For $F \in F_I$ define $F^{-1}(y)$ on $[0, 1]$ as $F^{-1}(y) \equiv \inf\{x \in R \mid F(x) \geq y\}$ where the infimum of the empty set is taken to be $+\infty$.

Remark 2.1. For $F, G \in F_I$ the following hold

- (i) $S(F)$ is a closed interval of R
- (ii) $I(F) \equiv \{x \in R \mid 0 < F(x) < 1\}$ is an open interval of R
- (iii) F is strictly increasing on the open interval $I(F)$
- (iv) $F^{-1}(\cdot)$ is continuous and strictly increasing on $(0, 1)$
- (v) $\phi(x) \equiv G^{-1}(F(x))$ is the unique strictly increasing function on $I(F)$ such that if X has cdf F then $\phi(X)$ has cdf G .

A weak ordering, $<$, on F_I is a relation on F_I which is reflexive and transitive. Every weak ordering $<$ on F_I can be "extended" to a partial order by first defining an equivalence relation \sim on F_I by $F \sim G \iff F < G$ and $G < F$. The resulting set of equivalence classes $\{E(F) \mid F \in F_I\}$ is partially ordered by the relation \leq which is well defined by $E(F) \leq E(G) \iff F < G$.

We now define a class of weak orderings which is a modification of a class proposed by Panchapakesan (1969).

Let H be a set of functions $h: I^2 \rightarrow I$ satisfying

$$h(x_1, x_2) \leq \max\{x_1, x_2\} \quad \forall (x_1, x_2) \in I^2.$$

Definition 2.1. For $F, G \in F_I$, F is said to be H -ordered wrt G (written $F \prec_H G$) $\Leftrightarrow G^{-1}F(h(x_1, x_2)) \leq h(G^{-1}F(x_1), G^{-1}F(x_2)) \forall h \in H$ and $x_1, x_2 \ni x_1, x_2$ and $h(x_1, x_2) \in I(F)$.

Lemma 2.1. H ordering is a weak ordering on F_I .

Proof. Reflexivity is straightforward. To show transitivity suppose

$F \prec_H G, G \prec_H J$ and $h \in H$ and x_1, x_2 satisfy x_1, x_2 and $h(x_1, x_2) \in I(F)$. Then $G^{-1}F(x_1), G^{-1}F(x_2)$ and $G^{-1}F(h(x_1, x_2)) \in I(G)$ and hence

$$(2.1) \quad \begin{aligned} G^{-1}Fh(x_1, x_2) &\leq h(G^{-1}F(x_1), G^{-1}F(x_2)) \\ &\leq \max\{G^{-1}F(x_1), G^{-1}F(x_2)\} \end{aligned}$$

implies $h(G^{-1}F(x_1), G^{-1}F(x_2)) \in I(G)$ by convexity. So

$$\begin{aligned} J^{-1}F(h(x_1, x_2)) &= J^{-1}GG^{-1}F(h(x_1, x_2)) \\ &\leq J^{-1}G(h(G^{-1}F(x_1), G^{-1}F(x_2))) \text{ by (2.1)} \\ &\leq h(J^{-1}F(x_1), J^{-1}F(x_2)) \text{ since } G \prec_H J \end{aligned}$$

and the proof is completed.

The motivation for introducing H -ordering is that it provides a convenient framework for embedding several well-known orderings.

Example 2.1. Let $I = [0, \infty)$ and H_1 consist of the single function

$$h(x_1, x_2) = \begin{cases} x_1 - x_2 & , x_1 > x_2 \\ 0 & , x_1 \leq x_2. \end{cases}$$

Now $h(x_1, x_2) \leq \max\{x_1, x_2\}$ for any $x_1, x_2 \geq 0$ and so H_1 is a weak ordering on F_0 . Furthermore

$$F \prec_{H_1} G \iff G^{-1}F(t-x) \leq G^{-1}F(t) - G^{-1}F(x) \text{ whenever } x, t, t-x \in I(F)$$

$$\iff G^{-1}F \text{ is superadditive on } I(F).$$

In particular $F \prec_{H_1} G(x) = 1 - e^{-x} \iff \bar{F}(t-x)\bar{F}(x) \geq \bar{F}(t)$ whenever x, t and $t-x \in I(F)$ where $\bar{F}(y) = 1 - F(y) \iff F$ is an absolutely continuous NBU (New Better than Used) distribution with convex support.

Before giving further examples recall the following definitions. For $F, G \in F_0$ (1) F is convex ordered wrt G ($F \prec_c G$) means $G^{-1}F(x)$ is convex on $I(F)$ and (2) F is star ordered wrt G ($F \prec_* G$) means $G^{-1}F(x)$ is starshaped on $I(F)$ i.e. $G^{-1}F(\lambda x) \leq \lambda G^{-1}F(x)$ whenever $x, \lambda x \in I(F)$ and $\lambda \in [0, 1]$. Convex ordering was introduced as an alternative to the standardized third moment inequality definition of skewness. Namely $F \prec_c G$ is interpreted to mean " G is more skewed to the right than F " (see Van Zwet (1964) pg. 9). Since the star shaped property can be thought of as a weakening of convexity (see Bruckner and Ostrow (1962)), $F \prec_* G$ can also be interpreted as an ordering according to skewness. For $F, G \in F$, F is said to be tail ordered wrt G ($F \prec_t G$) iff $G^{-1}F(x) - x$ is nondecreasing on $I(F)$. When $F \prec_t G$ then G is interpreted as having heavier tails than F (see Doksum (1969)).

Example 2.2. Take $I = [0, \infty)$ and $H_2 = \{h_\lambda \mid 0 \leq \lambda \leq 1\}$ where $h_\lambda(x_1, x_2) \equiv \lambda x_1 + (1-\lambda)x_2 \leq \max\{x_1, x_2\} \forall \lambda \in [0, 1]$; hence H_2 ordering is a weak ordering. It is straightforward to check that $F \prec_{H_2} G \iff G^{-1}F(\cdot)$ is

convex on $I(F)$. So H_2 yields convex ordering.

Example 2.3. Again let $I = [0, \infty)$ and define $H_3 = \{h_\lambda \mid 0 \leq \lambda \leq 1\}$ where $h_\lambda(x_1, x_2) = \lambda x_1 \leq \max\{x_1, x_2\} \forall 0 \leq \lambda \leq 1$ and so H_3 defines a weak ordering. It is straightforward to verify that H_3 yields star ordering.

Example 2.4. Choose $I = \mathbb{R}$ and $H_4 = \{h_\delta \mid 0 \leq \delta < \infty\}$ where $h_\delta(x_1, x_2) = x_1 - \delta \leq \max\{x_1, x_2\} \forall \delta \in [0, \infty)$; H_4 ordering is a weak ordering which reduces to tail ordering.

The remainder of this section is devoted to deriving some properties of star and tail ordering which will be used in the later sections.

Remark 2.1. It is straightforward to show that the set of distributions equivalent to $F \in F_0$ under star ordering ($F \in F$ under tail ordering) is $E_*(F) \equiv \{F(\delta x) \mid 0 < \delta < \infty\}$ ($E_t(F) \equiv \{F(x+\beta) \mid \beta \in \mathbb{R}\}$), the set of all scale changes (location shifts) of F .

Lemma 2.2.

- (a) For $F, G \in F_0$, $F \prec_* G \iff G^{-1}F(x)/x$ is nondecreasing on $R^+(F) \equiv \{x > 0 \mid F(x) < 1\}$.
- (b) For $F, G \in F$, $F \prec_t G \iff G^{-1}F(x) - x$ is nondecreasing on $R(F) \equiv \{x \in \mathbb{R} \mid F(x) < 1\}$.

Proof. Write $R^+(F)$ as the disjoint union $\{x > 0 \mid F(x) = 0\} \cup I(F)$ and $\forall x \in \{x > 0 \mid F(x) = 0\}$, $G^{-1}F(x)/x = G^{-1}(0)/x = -\infty$ and the result follows since $G^{-1}F(x)/x$ is nondecreasing on $I(F)$. The reverse implication is obvious and the proof for \prec_t is similar.

An important class of distributions preserving weak orderings are the

distributions of the order statistics. If $X_{(q)}$ ($1 \leq q \leq n$) is the q th order statistic based on X_1, \dots, X_n iid F then $P_F[X_{(q)} \leq x] = B(F(x); q, n)$ where

$$B(p; q, n) = \frac{n!}{(n-q)!(q-1)!} \int_0^p x^{q-1}(1-x)^{n-q} dx.$$

Theorem 2.1. If $F, G \in F_I$ and $F \prec_H G$ then $B(F; q, n) \prec_H B(G; q, n)$ holds $\forall 1 \leq q \leq n$. The proof is immediate from the definition of H ordering since $G^{-1}F(x) = G^{-1}B^{-1}BF(x)$ where $B(\cdot)$ and $B^{-1}(\cdot)$ denote $B(x; q, n)$ and its inverse respectively.

The final part of this section will describe stochastic bounds for several classes of distributions. For G in F_0 with $I(G) = (0, \infty)$ and $0 < \xi < \infty$ define

$$(2.2) \quad F_*(G) = \{F \in F_0 \mid F \prec_* G\}$$

$$(2.3) \quad F_*(G, \xi) = \{F \in F_*(G) \mid x_\alpha(F) = \xi\}.$$

Remark 2.3. $F_*(G) = F_*(G_\lambda) \forall \lambda > 0$ where $G_\lambda(x) \equiv G(\lambda x)$ since $\forall \lambda > 0$, $G_\lambda \in E_*(G)$ and $F \in F_*(G) \iff F \prec_* G \iff F \prec_* G_\lambda \iff F \in F_*(G_\lambda)$.

Example 2.5. Let $G(x) = 1 - e^{-x}$ or 0 as $x > 0$ or $x \leq 0$ respectively. Then $G^{-1}(y) = -\ln(1-y)$ for $0 \leq y < 1$ and $F \in F_*(G) \iff F$ is continuous on $[0, \infty)$ and $-\ln(1-F(x))/x$ is nondecreasing on $R^+(F) \iff F$ is a continuous IFRA distribution. In particular $F_*(G) = F_*(G_\lambda) \forall \lambda > 0$ where $G_\lambda(x) = 1 - e^{-\lambda x}$, $x > 0$.

Associate with $F_*(G)$ the families of distribution

$$(2.4) \quad F_{*,s}(x|\xi) = \begin{cases} G(xG^{-1}(\alpha)/\xi), & x < \xi \\ 1, & x \geq \xi \end{cases}$$

$$(2.5) \quad F_{*,l}(x|\xi) = \begin{cases} 0, & x < \xi \\ G(xG^{-1}(\alpha)/\xi), & x \geq \xi \end{cases}$$

where $0 < \xi < \infty$.

Remark 2.4. $G^{-1}F_{*,s}(\cdot|\xi)$ and $G^{-1}F_{*,l}(\cdot|\xi)$ are starshaped on $R^+(F_{*,s}(\cdot|\xi))$ and $R^+(F_{*,l}(\cdot|\xi))$ respectively and $x_\alpha(F_{*,s}(\cdot|\xi)) = x_\alpha(F_{*,l}(\cdot|\xi)) = \xi$. However neither $F_{*,s}(\cdot|\xi)$ nor $F_{*,l}(\cdot|\xi)$ is in F_0 and consequently neither is in $F_*(G)$.

Theorem 2.2. For all $0 < \xi < \infty$ and $F \in F_*(G, \xi)$,

$$F_{*,s}(\cdot|\xi) \prec_{st} F \prec_{st} F_{*,l}(\cdot|\xi).$$

Proof. If $R^+(F) = (0, M)$ then $x \leq 0 \Rightarrow F(x) = G(x) = F_{*,s}(x|\xi) = F_{*,l}(x|\xi) = 0$ and $\forall 0 < x < \xi \leq y < M$

$$\begin{aligned} G^{-1}F(x)/x &\leq G^{-1}F(\xi)/\xi = G^{-1}(\alpha)/\xi \leq G^{-1}F(y)/y \\ \Leftrightarrow F(x) &\leq G(xG^{-1}(\alpha)/\xi) \leq G(yG^{-1}(\alpha)/\xi) \leq F(y). \end{aligned}$$

If $M = +\infty$ the proof is complete while if $M < +\infty$ then $\forall y \geq M$ $F(y) = 1 \geq G(yG^{-1}(\alpha)/\xi)$ and the result follows. See Figure 2.1.

Since both $\{F_{*,s}(\cdot|\xi) | 0 < \xi < \infty\}$ and $\{F_{*,l}(\cdot|\xi) | 0 < \xi < \infty\}$ are (stochastically increasing) scale parameter families the result can be slightly strengthened as follows.

Corollary 2.1. For $0 < \xi' \leq \xi \leq \xi''$

- (a) $F \in F_*(G, \xi') \Rightarrow F \prec_{st} F_{*,l}(\cdot|\xi)$ and
- (b) $F \in F_*(G, \xi'') \Rightarrow F_{*,s}(\cdot|\xi) \prec_{st} F$

Remark 2.5. It can be shown that the bounds $F_{*,s}(\cdot|\xi)$ and $F_{*,l}(\cdot|\xi)$

are tight in the sense that they are the stochastically largest and smallest distributions respectively satisfying Corollary 2.1.

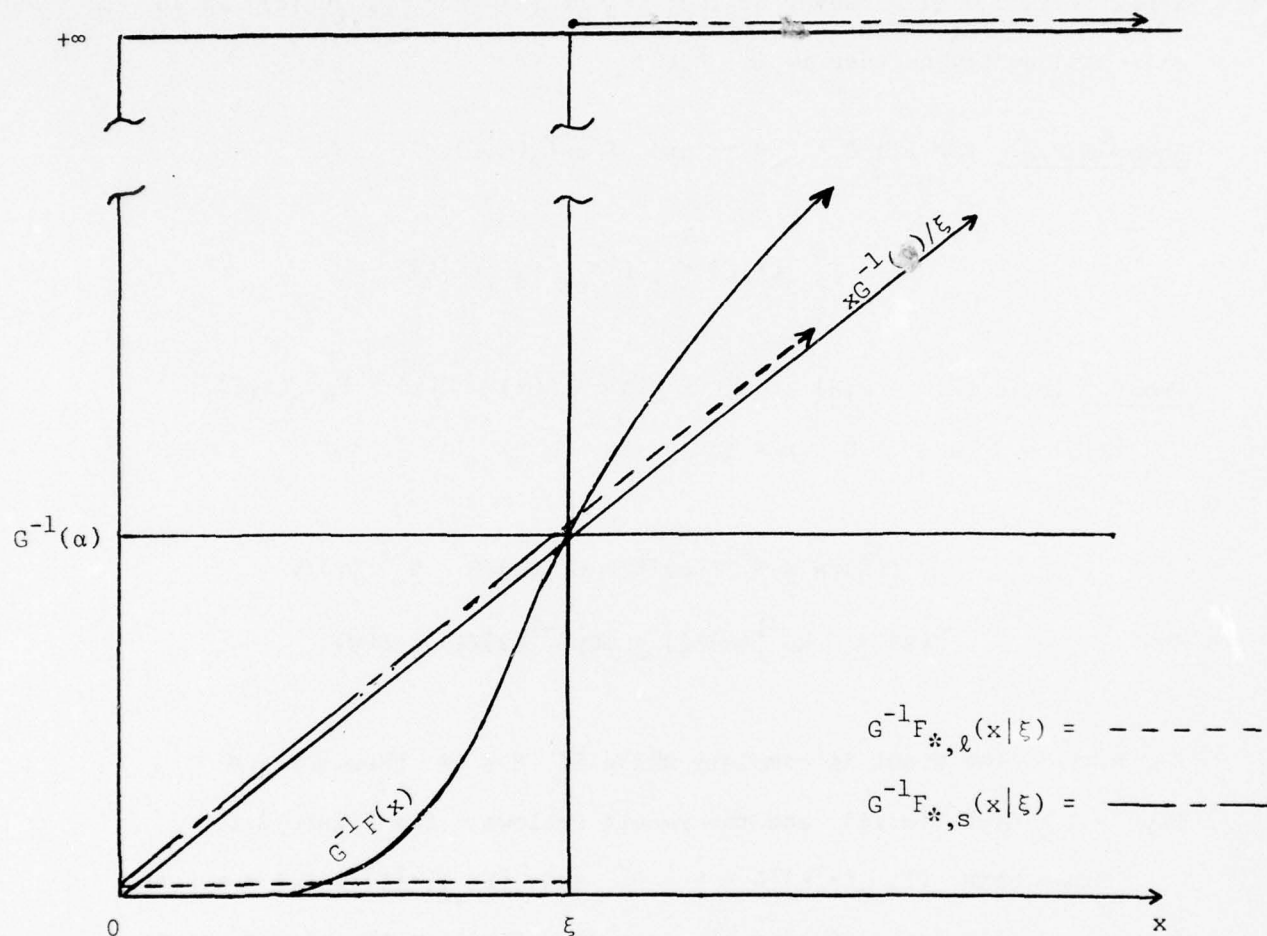


Figure 2.1

The analogous results for tail ordered families are as follows. For $G \in \mathcal{F}$ with $I(G) = R$ and $\xi \in R$ define

$$F_t(G) = \{F \in \mathcal{F} | F \prec_t G\}$$

$$F_t(G, \xi) = \{F \in F_t(G) \mid x_\alpha(F) = \xi\}$$

$$F_{t,s}(x|\xi) = \begin{cases} G(x + G^{-1}(\alpha) - \xi) & , x < \xi \\ 1 & , x \geq \xi \end{cases}$$

$$F_{t,l}(x|\xi) = \begin{cases} 0 & , x < \xi \\ G(x + G^{-1}(\alpha) - \xi) & , x \geq \xi. \end{cases}$$

Example 2.6. When $G(x) = 1/(1+e^{-x})$, $x \in \mathbb{R}$ then $G^{-1}(y) = -\ln((1-y)/y)$ on $(0,1)$

$$\Rightarrow F_t(G) = \left\{ F \in F \mid \frac{\bar{F}(t+x)/F(t+x)}{\bar{F}(t)/F(t)} \leq e^{-x} \forall x \geq 0 \text{ and } t \in I(F) \right\} \\ (t+x) \in I(F)$$

where $\bar{F}(y) = 1 - F(y)$ is the reliability function of F . In words $F_t(G)$ is the set of all continuous distributions such that the ratio of the odds of surviving beyond $t+x$ to the odds of surviving beyond t is uniformly bounded above (in t) by e^{-x} .

Remark 2.6. $G^{-1}F_{t,s}(x|\xi) - x$ and $G^{-1}F_{t,l}(x|\xi) - x$ are nondecreasing on $\{x \in \mathbb{R} \mid F_{t,s}(x|\xi) < 1\}$ and $\{x \in \mathbb{R} \mid F_{t,l}(x|\xi) < 1\}$ respectively and $x_\alpha(F_{t,s}(\cdot|\xi)) = x_\alpha(F_{t,l}(\cdot|\xi)) = \xi$ but neither is in F and hence not $F_t(G)$. Both $\{F_{t,s}(\cdot|\xi) \mid \xi \in \mathbb{R}\}$ and $\{F_{t,l}(\cdot|\xi) \mid \xi \in \mathbb{R}\}$ are (stochastically increasing) locations parameter families.

Theorem 2.3. For all $\xi \in \mathbb{R}$ and $F \in F_t(G, \xi)$, $F_{t,s}(\cdot|\xi) \leq_{st} F \leq_{st} F_{t,l}(\cdot|\xi)$. The proof is similar to that of Theorem 2.3.

Corollary 2.2. For $\xi' \leq \xi \leq \xi''$

$$(a) \quad F \in F_t(G, \xi') \Rightarrow F \prec_{st} F_{t, \ell}(\cdot | \xi)$$

$$(b) \quad F \in F_t(G, \xi'') \Rightarrow F_{t, s}(\cdot | \xi) \prec_{st} F.$$

Finally it should be noted that $F_{t, s}(\cdot | \xi)$ and $F_{t, \ell}(\cdot | \xi)$ are the stochastically largest and smallest distributions respectively satisfying Corollary 2.2.

3. Fixed Size Subset Selection Procedures

3.1 Preferred Population Formulation

When faced with the problem of selecting at least r of the t best of k populations ($1 \leq r \leq t \leq k$) an experimenter may well be indifferent between the selection of one of the t best populations and one of the $(k-t)$ worst populations when they are sufficiently close together. This is the rationale of the indifference zone formulation (IZF) of selection problems in which the procedure is only required to guarantee correct selection (CS) with prespecified probability, P^* , when the t^{th} and $(t+1)^{\text{st}}$ best populations are sufficiently "far" apart (see Mahamunulu (1967)).

Alternatively, if the t^{th} best and one of the $(k-t)$ worst populations are sufficiently "close" so that an experimenter would be satisfied with the selection of either one then one can define a preferred population as any population satisfying this criterion and a correct selection as the selection of at least r preferred populations. The preferred population formulation (PPF) requires the procedure make a correct selection with at least probability P^* no matter what the true vector F_{α} of distributions.

Suppose $x_{\alpha}(F_i) \in E$ where E is a known interval of the real line. Then following Santner (1975) consider measures of closeness to $\pi_{(k-t+1)}$ defined in terms of a function $p: E \rightarrow R$ satisfying:

$$(3.1) \quad p(\cdot) \text{ is continuous}$$

$$(3.2) \quad p(\xi) \geq \gamma \equiv \inf\{\xi \in E\} \quad \forall \xi \in E$$

$$(3.3) \quad p(\cdot) \text{ is strictly increasing on } E' \equiv \{\xi \in E \mid p(\xi) > \gamma\}$$

$$(3.4) \quad p(\xi) < \gamma \quad \forall \xi \in E'.$$

Definition 3.1. π_i is a preferred population (relative to $p(\cdot)$) iff $x_{\alpha}(F_i) > p(x_{\alpha}(F_{[k-t+1]}))$.

Clearly at least $\pi_{(k-t+1)}, \dots, \pi_{(k)}$ are preferred populations and the true population configuration will contain between t and k preferred π_i 's.

Remark 3.1. The PPF yields a strengthening of the probability guarantee over the IZF. Consider the problem of selecting the $t(r=t)$ best π_i 's when the true $F \in \Omega \equiv F_I^k$ where $\Omega' = \{F \in \Omega \mid x_\alpha(F_{[k-t]}) > p(x_\alpha(F_{[k-t+1]}))\}$ and $\bar{\Omega}' = \Omega \setminus \Omega'$ are the indifference zone and preference zone respectively.

The corresponding problem based on the PPF is to select at least t preferred populations (relative to $p(\cdot)$). For $F \in \bar{\Omega}'$ the t best populations are the only preferred π_i 's and both the IZF and PPF guarantee their selection with probability P^* . For $F \in \Omega'$ the PPF still guarantees the selection of at least t preferred π_i 's while the IZF guarantees nothing.

Two specific choices for $p(\cdot)$ are now derived. Recall that for $F \in F_*(G)$ any scale change of $F(x)$, say $F(\delta x)$, is equivalent to $F(x)$. Hence we require that π_i be preferred relative to $\pi_{(k-t+1)}$ independent of the selected measurement units. More precisely this means π_i is close to $\pi_{(k-t+1)}$ when $F_i \in F_*(G, \xi'/\delta)$ and $F_{[k-t+1]} \in F_*(G, \xi/\delta)$ independent of $\delta > 0$, i.e. $p(\cdot)$ must satisfy $\forall \xi', \xi \in (0, \infty)$

$$(3.5) \quad \xi' > p(\xi) \iff \xi'/\delta > p(\xi/\delta) \quad \forall \delta > 0.$$

It can be shown that the unique $p(\cdot)$ satisfying (3.1) - (3.5) is $p(\xi) = c^*\xi$ where $0 < c^* < 1$. In the tail ordered case a similar argument based on the location invariant classes $E_t(G)$ leads to the choice $p(\xi) = \xi - d^*$ where $0 < d^* < \infty$. Hence π_i will be called preferred in the star (tail) ordered case iff $x_\alpha(F_i) > c^*x_\alpha(F_{[k-t+1]}) (x_\alpha(F_{[k-t+1]}) - d^*)$.

In the remainder of the paper G is a fixed distribution in $F_0(F)$ with $I(G) = (0, \infty)$ (R) in discussions of the star (tail) ordered model. Let

$$(3.6) \quad \Omega_* = \{F = (F_1, \dots, F_k) \mid F_i \in F_*(G) \forall i\} \text{ and}$$

$$(3.7) \quad \Omega_t = \{F = (F_1, \dots, F_k) \mid F_i \in F_t(G) \forall i\}.$$

Given $t(1 \leq t \leq k)$, $r(1 \leq r \leq t)$, $s(r \leq s \leq k-t+r)$ and $\alpha(0 < \alpha < 1)$ the goal is to select a subset of size s containing at least r preferred populations (a correct section). Note that for $k-t+r \leq s \leq k$ a correct selection must automatically occur. Let $N_\alpha \equiv \min\{n \geq 1 \mid 1 \leq (n+1)\alpha \leq n\}$ and for each $n \geq N_\alpha$ define the procedure.

$R_1(s, n)$: Take independent random samples of size n from each population and select $\pi_i \iff T_i \geq T_{[k-s+1]}$ where $T_{[1]} \leq \dots \leq T_{[k]}$ are the ordered sample α -quantiles T_1, \dots, T_k .

Design Requirement: Given $P^* \in \left(\sum_{j=r}^{\min\{t, s\}} \binom{t}{j} \binom{k-t}{s-j} / \binom{k}{s}, 1 \right)$ and $0 < c^* < 1$ ($0 < d^* < \infty$) determine the smallest $n \geq N_\alpha$ so that

$$(3.8) \quad P_F[CS \mid R_1(s, n)] \geq P^* \quad \forall F \in \Omega_*(\Omega_t)$$

where the event $[CS \mid R_1(s, n)]$ denotes a correct selection using $R_1(s, n)$.

3.2 Selection from Star Ordered Families

For fixed $\alpha \in (0, 1)$ it follows without loss of generality from Remark 2.3 that G can be chosen so that $G^{-1}(\alpha) = 1$. Define the following subsets of Ω_* . For $0 < \xi < \infty$ and $t \leq i \leq k$ let

$$(3.9) \quad \Omega_*(i) = \begin{cases} \{F \in \Omega_* \mid x_\alpha(F_{[k-1]}) \leq c^* x_\alpha(F_{[k-t+1]}) < x_\alpha(F_{[k-i+1]})\}, & t \leq i < k \\ \{F \in \Omega_* \mid c^* x_\alpha(F_{[k-t+1]}) < x_\alpha(F_{[1]})\} & , i=k, \end{cases}$$

$$(3.10) \quad \Omega_{*,\xi} = \{F \in \Omega_* \mid x_\alpha(F_{[k-t+1]}) \approx \xi\},$$

$$(3.11) \quad \Omega_*(i, \xi) = \Omega_*(i) \cap \Omega_{*,\xi}.$$

It follows that $\Omega_* = \bigcup_{\xi > 0} \Omega_{*,\xi} = \bigcup_{\xi > 0} \bigcup_{i=t}^k \Omega_*(i, \xi)$ and, for example, $\Omega_*(i, \xi)$ is the set of configurations for which there are exactly i preferred populations and the t^{th} smallest α -quantile is ξ .

It is easy to see that for $F \in \Omega_*(i)$ where $r + k - s \leq i \leq k$ that $P_F[CS|R_1(s, n)] = 1$. Hence $\inf_{\Omega_*} P_F[CS|R_1(s, n)] = \min_{i \in I} \inf_{F \in \Omega_*(i)} P_F[CS|R_1(s, n)]$ where $I = \{i \mid t \leq i < r + k - s\}$. For $F \in \Omega_*$ and $i \in I$ define

$T_{(i)} \equiv$ sample α -quantile from $\pi(i)$,

$W_i \equiv r^{\text{th}}$ largest of $\{T_{(k-i+1)}, \dots, T_{(k)}\}$ and

$A_i \equiv \{W_i \geq \text{at least } (r+k-s-i) \text{ of } T_{(1)}, \dots, T_{(k-i)}\}.$

Remark 3.2. It can be shown that the A_j 's form a nondecreasing sequence of events and that $\chi_{A_i}(T)$, the characteristic function of A_i , is nonincreasing in any of $T_{(1)}, \dots, T_{(k-i)}$ and is nondecreasing in any of $T_{(k-i+1)}, \dots, T_{(k)}$. This latter property and an application of Theorem 2.1 implies that $P_F[A_i] = E_F[\chi_{A_i}(T)]$ is nonincreasing in any of $F_{[1]}, \dots, F_{[k-i]}$ and is nondecreasing in any of $F_{[k-i+1]}, \dots, F_{[k]}$ under stochastic ordering

\prec_{st} (see Mahumunulu (1967) or Alam and Rizi (1966)).

Lemma 3.1. $\inf_{\Omega_{*}(i)} P_{\tilde{F}}[CS|R_1(s,n)] = \inf_{\Omega_{*}(i,1)} P_{\tilde{F}}[A_i]$ for $i \in I$.

Proof. Pick $\tilde{F} \in \Omega_{*}(i, \xi)$ then $P_{\tilde{F}}[CS|R_1(s,n)] = P_{\tilde{F}}[A_i] = P_{\tilde{F}}[W_i/\xi \geq \text{at least } (r+k-s-i) \text{ of } T_{(1)}/\xi, \dots, T_{(k-i)}/\xi]$. Now $T' = T/\xi$ has the vector of cdf's $\tilde{F}'(x) = (F_1(\xi x), \dots, F_k(\xi x))$ which is in $\Omega_{*}(i,1)$ by the equivalence of F_i under scale changes and the definition of preferred π_i . So $P_{\tilde{F}}[A_i] = P_{\tilde{F}'}[A_i]$ and the proof is complete.

Remark 3.3. An equivalent way of viewing the proof of Lemma 3.1 is as an application of invariance. The problem is invariant under the group of all possible common scale changes of the raw data. Hence the risk under a 0-1 loss structure becomes the probability of incorrect selection and is constant over orbits - the orbit of \tilde{F} consisting of all common scale changes of the components of \tilde{F} .

Next a lower bound for $P_{\tilde{F}}[CS|R_1(s,n)]$ over $\tilde{F} \in \Omega_{*,1}$ is obtained. Given $i \in I$ and $\tilde{F} \in \Omega_{*}(i,1)$ it follows from Corollary 2.1 that

$$(3.12) \quad \left\{ \begin{array}{ll} F_{[j]} \prec_{st} F_{*,\ell}(\cdot|c^*) & , 1 \leq j \leq k-i \\ F_{*,s}(\cdot|1) \prec_{st} F_{[j]} & , k-t+1 \leq j \leq k \\ F_{*,s}(\cdot|c^*) \prec_{st} F_{[j]} & , \text{for any other } j \end{array} \right.$$

since $x_{\alpha}(F_{[j]}) \leq c^*$ for $1 \leq j \leq k-i$, $1 \leq x_{\alpha}(F_{[j]})$ for $k-t+1 \leq j \leq k$ and $c^* < x_{\alpha}(F_{[j]})$ for all remaining j 's. Let $\tilde{F}(i)$ be any configuration having $k-i$, t and $i-t$ (if $i > t$) components of the types $F_{*,\ell}(\cdot|c^*)$, $F_{*,s}(\cdot|1)$ and $F_{*,s}(\cdot|c^*)$ respectively. Then

$$\begin{aligned}
P_{\tilde{F}}[LCS|R_1(s,n)] &= P_{\tilde{F}}[A_i] \\
&\geq P_{\tilde{F}(i)}[A_i] \quad \text{from (3.12) and Remark 3.2} \\
&\geq P_{\tilde{F}(i)}[A_t] \quad \text{since } A_t \subset A_i \\
&\geq P_{\tilde{F}(t)}[A_t]
\end{aligned}$$

where the final inequality holds since when $i > t$ the $j^{\text{th}}(k-i+1 \leq j \leq k-t)$ ordered component of $\tilde{F}(i)$ is $F_{*,s}(\cdot|c^*) \prec_{st} F_{*,\ell}(\cdot|c^*)$, the j^{th} ordered component of $\tilde{F}(t)$.

To calculate the lower bound $P_{\tilde{F}(t)}[A_t]$ let

$$(3.13) \quad \begin{cases} H(y) = B(B(y); k-t-s+r, k-t) \\ J(y) = B(B(y); t-r+1, t) \end{cases}$$

where $B(y) = B(y; q, n)$, $q = [(n+1)\alpha]$ and $[\cdot]$ denotes the greatest integer function. Hence $H(F_{*,\ell}(y|c^*))$ is the cdf of the $(k-t-s+r)^{\text{th}}$ smallest of $\{T_{(1)}, \dots, T_{(k-t)}\}$ and $J(F_{*,s}(y|1))$ is the cdf of the $(t-r+1)^{\text{th}}$ smallest of $\{T_{(k-t+1)}, \dots, T_{(k)}\}$ all under $\tilde{F}(t)$.

Therefore

$$\begin{aligned}
P_{\tilde{F}(t)}[A_t] &= \int_0^\infty H(F_{*,\ell}(y|c^*)) dJ(F_{*,s}(y|1)) \\
&= \int_{c^*}^1 H(G(y/c^*)) dJ(G(y)) \\
&\quad + H(G(1/c^*))(1 - J(\alpha)) \\
&= \int_{G(c^*)}^\alpha H(G(G^{-1}(y)/c^*)) dJ(y) \\
&\quad + H(G(1/c^*))(1 - J(\alpha)) \\
&= F_n^*(r, t, s, k, \alpha, c^*) \quad \text{or simply } F_n^*.
\end{aligned}$$

Since $F(t) \notin \Omega_*$ (the component distributions are not in $F_*(G)$) it is not immediate that F_n^* is the infimum of the PCS over Ω_* . The next result establishes that this is the case.

Theorem 3.1. $\inf_{\Omega_*} P_{F_j}[CS|R_1(s,n)] = F_n^*(r,t,s,k,a,c^*)$

Proof. From Lemma 3.1 and the discussion above it suffices to exhibit a sequence of configurations $\{F_j\}$ in Ω_* so that $\lim_{j \rightarrow \infty} P_{F_j}[CS|R_1(s,n)] = F_n^*$. A sequence in Ω_* which approximates $F(t)$ is the obvious candidate. Let $L \equiv [1/c^*] + 1$ and for $j \geq L$ define

$$F_{\ell,j}(y) \begin{cases} 0 & , y < c^* - j^{-1} \\ G[j(y - (c^* - j^{-1}))] & , c^* - j^{-1} \leq y < c^* \\ G(y/c^*) & , c^* \leq y \end{cases}$$

$$F_{s,j}(y) \begin{cases} G(y) & , y < 1 \\ G(1+j(y-1)) & , 1 \leq y. \end{cases}$$

Then $F_{\ell,j}(\cdot) \xrightarrow{w} F_{*,\ell}(\cdot|c^*) \Rightarrow H(F_{\ell,j}(\cdot)) \xrightarrow{w} H(F_{*,\ell}(\cdot|c^*))$ and

$F_{s,j}(\cdot) \xrightarrow{w} F_{*,s}(\cdot|1) \Rightarrow J(F_{s,j}(\cdot)) \xrightarrow{w} J(F_{*,s}(\cdot|1))$ where \xrightarrow{w} denotes weak

convergence of distributions. Furthermore $F_{\ell,j} \in F_t(G,c^*)$ and

$F_{s,j} \in F_*(G,1)$ and so F_j with t components $F_{\ell,j}$ and $(k-t)$ components $F_{s,j}$ is in $\Omega_*(t)$. Therefore

$$\begin{aligned}
P_{F_{\sim j}}[A_*] &= \int_0^\infty H(F_{\ell,j}(y)) dJ(F_{s,j}(y)) \\
&= \int_0^{c^*} H(F_{\ell,j}(y)) dJ(F_{*,s}(y|1)) + \int_{c^*}^\infty H(F_{*,\ell}(y|c^*)) dJ(F_{s,j}(y))
\end{aligned}$$

since $F_{\ell,j}(y) = F_{*,\ell}(y|c^*)$ on $[c^*, \infty)$ and

$$F_{s,j}(y) = F_{*,s}(y|1) \text{ on } [0,1]$$

$$= \int_0^{c^*} H(F_{\ell,j}(y)) dJ(F_{*,s}(y|1)) + \int_0^\infty H(F_{*,\ell}(y|c^*)) dJ(F_{s,j}(y))$$

since $F_{*,\ell}(y|c^*) = 0$ on $[0, c^*)$

$$\rightarrow 0 + \int_0^\infty H(F_{*,\ell}(y|c^*)) dJ(F_{*,s}(y|1)) = F_n^*$$

where the first convergence follows from dominated convergence since $H(F_{\ell,j}(y)) \rightarrow 0$ a.e. $[J(F_{*,s}(\cdot|1))]$ on $[0, c^*]$ and the second follows from the weak convergence of $J(F_{s,j}(\cdot))$ to $J(F_{*,s}(\cdot|1))$ and the fact that $H(F_{*,\ell}(\cdot|c^*))$ is bounded and continuous a.e. $[J(F_{*,s}(\cdot|1))]$ on $[0, \infty)$. The proof is completed.

It can be shown that $F_n^* \rightarrow 1$ as $n \rightarrow \infty$ and hence (3.8) can be guaranteed by a finite n for any $P^* < 1$.

Example 3.1. Given $\alpha \in (0,1)$ choose $G(x) = \begin{cases} 1 - e^{x \ln(1-\alpha)}, & x \geq 0; \\ 0, & x < 0 \end{cases}$

Ω_* is then all configurations of continuous IFRA distributions. Note that $G(\cdot)$ is chosen so that $G^{-1}(\alpha) = 1$. Suppose it is desired to select at least $r = t$ preferred populations. In this case

$$F_n^* = t \int_{1-(1-\alpha)^{1/c^*}}^{\alpha} B(B(1 - [1-x]^{1/c^*}); k-s, k-t) dB(x) \\ + B(B(1-[1-\alpha]^{1/c^*}); k-s, k-t)(1 - B(\alpha))^t$$

For computational purposes it is desirable to remove the composition of incomplete beta functions from inside the integral. After an integration by parts and a change of variables F_n^* becomes

$$(3.14) \quad \binom{k-t}{k-s} \int_{\alpha}^{1-(1-\alpha)^{1/c^*}} (1 - B(1 - [1-y]^{c^*}))^t (1 - B(y))^{s-t} dB^{k-s}(y) \\ + (1 - B(1 - (1-\alpha)^{c^*}))^t B(B(\alpha); k-s, k-t)$$

Expression (3.14) was evaluated on Cornell University's IBM 370/169 computer using a procedure based on Simpson's rule (see Shampine and Allen (1973)). Tables 1,2 and 3 give the smallest odd sample sizes required to meet (3.8) for $\alpha = .25(.25).75$, $c^* = .65, .70, p^* = .75, .90, .95$, $k = 2(1)9$, $t = 1(1)[k/2]$, $r = t$, $s = t(1)[k/2]$.

One measure of the cost of using the nonparametric procedure $R_1(s,n)$ is obtained by comparing the sample size required by $R_1(s,n)$ to achieve a specified p^* to that required by the best procedure for the parametric problem with exponential populations. Suppose π_i has cdf $F(x|\theta_i) = 1 - e^{-x/\theta_i}$, $x > 0$ where $\theta_i > 0$ but unknown then $x_{\alpha}(F_i) = -\theta_i \ln(1-\alpha) \Rightarrow x_{\alpha}(F(\cdot|\theta_i)) > c^* x_{\alpha}(F_{[k-t+]}) \Leftrightarrow \theta_i > c^* \theta_{[k-t+1]}$ where $\theta_{[i]}$ is the i^{th} ordered θ_j . In this case the optimal selection procedure is the natural procedure based on sample means rather sample

k	t	s	P*		
			.75	.90	.95
2	1	1	39	87	131
3	1	1	55	115	163
4	1	1	69	135	187
4	1	2	27	67	103
4	2	2	87	155	211
5	1	1	79	147	203
5	1	2	35	79	119
5	2	2	103	179	235
6	1	1	89	159	215
6	1	2	43	91	131
6	1	3	23	59	95
6	2	2	119	195	255
6	2	3	59	115	155
6	3	3	127	207	267
7	1	1	97	169	191
7	1	2	47	99	143
7	1	3	27	67	103
7	2	2	127	211	271
7	2	3	71	127	171
7	3	3	143	227	287
8	1	1	103	177	235
8	1	2	53	107	151
8	1	3	31	75	111
8	1	4	23	59	87
8	2	2	139	219	283
8	2	3	79	139	183
8	2	4	55	99	139
8	3	3	155	239	299
8	3	4	91	151	195
8	4	4	163	247	307
9	1	1	109	183	243
9	1	2	59	115	163
9	1	3	35	83	123
9	1	4	23	63	95
9	2	2	147	231	291
9	2	3	87	151	195
9	2	4	59	111	147
9	3	3	167	251	313
9	3	4	99	163	211
9	4	4	175	263	325

Table 3.1. Sample sizes required by $R_1(s, n)$ when selecting from Ω_* when $G(x) = 1 - e^{-x}$ with $\alpha = .25$ and $c^* = .65$.

			$c^* = .65$			$c^* = .70$		
			P^*			P^*		
k	t	s	.75	.90	.95	.75	.90	.95
2	1	1	19	45	67	27	65	97
3	1	1	29	59	85	43	85	123
4	1	1	37	69	97	53	101	141
4	1	2	13	33	51	19	47	73
4	2	2	43	79	109	63	117	157
5	1	1	43	77	105	61	113	153
5	1	2	17	39	59	25	57	85
5	2	2	53	91	121	77	133	177
6	1	1	47	83	111	69	121	163
6	1	2	21	45	67	31	67	97
6	1	3	11	29	47	17	43	67
6	2	2	61	101	131	87	147	191
6	2	3	31	57	79	43	83	115
6	3	3	65	107	137	95	155	199
7	1	1	51	87	117	75	129	171
7	1	2	25	51	73	37	73	105
7	1	3	13	35	51	19	49	75
7	2	2	67	107	139	97	155	203
7	2	3	35	65	87	51	95	127
7	3	3	73	115	147	107	169	215
8	1	1	53	91	121	79	135	179
8	1	2	27	51	79	41	81	113
8	1	3	17	39	57	23	55	83
8	1	4	11	29	45	15	41	63
8	2	2	71	113	145	103	165	211
8	2	3	41	71	95	59	103	137
8	2	4	27	51	69	37	73	101
8	3	3	79	123	155	117	179	225
8	3	4	47	77	101	67	111	145
8	4	4	83	127	159	121	185	231
9	1	1	57	95	125	83	139	185
9	1	2	31	59	83	45	87	121
9	1	3	19	43	61	27	61	89
9	1	4	11	31	49	17	45	69
9	2	2	75	119	149	109	173	219
9	2	3	45	77	101	65	111	145
9	2	4	29	55	75	43	81	109
9	3	3	85	129	161	125	189	235
9	3	4	51	83	107	75	121	155
9	4	4	89	135	167	131	193	243

Table 3.2. Sample sizes required by $R_1(s, n)$ when selecting from Ω^* when $G(x) = 1 - e^{-x}$ with $\alpha = .50$ and $c^* = .65, .70$.

k t s			c* = .65			c* = .70		
			p*			p*		
			.75	.90	.95	.75	.90	.95
2	1	1	15	33	51	21	47	73
3	1	1	21	45	63	33	65	93
4	1	1	29	53	73	41	77	105
4	1	2	11	27	39	15	35	55
4	2	2	33	59	81	47	87	119
5	1	1	33	57	79	45	85	115
5	1	2	13	31	43	19	43	63
5	2	2	41	69	91	59	101	133
6	1	1	37	63	85	53	91	123
6	1	2	17	35	51	23	51	71
6	1	3	11	23	35	11	31	51
6	2	2	45	75	99	67	111	143
6	2	3	23	43	59	35	63	85
6	3	3	49	79	103	71	115	149
7	1	1	37	67	89	57	97	129
7	1	2	19	39	55	27	55	79
7	1	3	11	27	39	15	35	55
7	2	2	49	81	105	73	119	151
7	2	3	27	49	65	39	71	95
7	3	3	55	87	111	81	127	161
8	1	1	41	69	93	61	101	135
8	1	2	21	41	59	29	61	85
8	1	3	13	29	43	19	41	61
8	1	4	7	23	31	11	31	47
8	2	2	53	85	109	79	125	159
8	2	3	31	53	71	43	77	103
8	2	4	19	39	51	27	55	75
8	3	3	61	93	117	87	135	169
8	3	4	35	57	75	51	83	107
8	4	4	63	95	119	91	139	173
9	1	1	45	73	95	65	105	139
9	1	2	23	45	63	33	65	91
9	1	3	15	31	47	21	45	67
9	1	4	11	23	35	15	35	51
9	2	2	57	89	113	83	129	165
9	2	3	33	57	75	49	83	109
9	2	4	23	43	57	31	59	81
9	3	3	65	97	121	93	141	177
9	3	4	39	63	79	55	91	115
9	4	4	67	101	125	99	147	181

Table 3.3. Sample sizes required by $R_1(s, n)$ when selecting from Ω_* when $G(x) = 1 - e^{-x}$ with $\alpha = .75$ and $c^* = .65, .70$.

α -quantiles (Eaton (1967)). Denote this procedure as $R_M(s,n)$. It is easy seen that for $\Omega = (0,\infty)^k$

$$\inf_{\Omega} P_{\theta} \{CS | R_M(s,n)\} = \int_0^{\infty} H(\Gamma(y|c^*,n)) dJ(\Gamma(y|1,n))$$

$$\text{where } \Gamma(y|\theta,n) = \int_0^y \frac{e^{-w/\theta} w^{n-1}}{(n-1)! \theta^n} dw.$$

when $r = t$ this integral reduces, after a change of variables, to

$$(3.15) \quad (k-s) \binom{k-t}{s-t} \int_0^1 \{1 - \Gamma(c^* \Gamma^{-1}[x|1,n]|1,n)\}^t x^{k-s-1} (1-x)^{s-t} dx.$$

Expression (3.15) was evaluated using an algorithm based on Simpson's rule. Table 4 contains the smallest odd sample signs required to satisfy

$$(3.16) \quad \inf_{\Omega} P_{\theta} [CS | R_M(s,n)] \geq P^*$$

for $c^* = .65, .70, P^* = .75, .90, .95, k = 2(1)9, r = t = 1(1)[k/2]$ and $s = t(1)[k/2]$.

These calculations show that the exponential procedure requires roughly half as many observations as the non parametric IFRA procedure.

3.3 Selection from Tail Ordered Families

Fix $G' \in F$ satisfying $I(G') = R$ and $\alpha \in (0,1)$. Choose $G \in E_t(G')$ so that $G^{-1}(\alpha) = 0$ and for $d^* \in (0,\infty)$ define

k	t	s	$c^* = .65$			$c^* = .70$		
			P^*			P^*		
			.75	.90	.95	.75	.90	.95
2	1	1	7	19	31	9	27	43
3	1	1	11	27	41	17	39	59
4	1	1	15	33	47	23	47	67
4	1	2	5	15	23	7	21	33
4	2	2	21	39	53	29	57	77
5	1	1	19	37	51	27	53	73
5	1	2	7	19	29	11	27	41
5	2	2	25	45	59	37	65	85
6	1	1	21	39	55	31	57	79
6	1	2	9	23	33	13	33	47
6	1	3	5	13	21	7	19	31
6	2	2	29	49	63	43	71	93
6	2	3	15	29	39	21	39	55
6	3	3	33	53	67	47	75	97
7	1	1	23	43	57	33	61	83
7	1	2	11	25	35	17	37	51
7	1	3	7	17	25	9	23	35
7	2	2	33	53	67	47	77	97
7	2	3	17	33	43	25	45	61
7	3	3	37	57	71	53	81	103
8	1	1	25	45	59	35	65	85
8	1	2	13	27	39	19	39	55
8	1	3	7	19	29	11	27	41
8	1	4	5	13	21	5	19	29
8	2	2	35	55	71	51	81	103
8	2	3	21	35	47	29	51	67
8	2	4	13	25	33	17	35	49
8	3	3	39	61	75	57	87	109
8	3	4	23	37	49	33	55	71
8	4	4	41	61	77	59	89	111
9	1	1	25	47	61	37	67	89
9	1	2	15	29	41	21	43	59
9	1	3	9	21	31	13	29	43
9	1	4	5	15	23	7	21	33
9	2	2	37	57	73	53	83	105
9	2	3	23	37	49	31	55	71
9	2	4	15	27	37	21	39	53
9	3	3	41	63	79	61	91	113
9	3	4	25	41	53	37	59	75
9	4	4	45	65	81	63	95	117

Table 3.4. Sample sizes required by $R_M(s,n)$ when selecting from exponential populations with α arbitrary and $c^* = .65, .70$.

$$(3.17) \quad \Omega_t \equiv \{F \in F^k \mid F_i \leq_t G \forall 1 \leq i \leq k\},$$

$$(3.18) \quad \Omega_{t,\xi} \equiv \{F \in \Omega_t \mid x_\alpha(F_{[k-t+1]}) = \xi\}, \xi \in R,$$

$$(3.19) \quad \Omega_t(i) \equiv \begin{cases} \{F \in \Omega_t \mid x_\alpha(F_{[k-i]}) \leq x_\alpha(F_{[k-t+1]}) - d^* < x_\alpha(F_{[k-i+1]})\}, & t \leq i < k \\ \{F \in \Omega_t \mid x_\alpha(F_{[k-t+1]}) - d^* < x_\alpha(F_{[1]})\} & , i=k, \end{cases}$$

$$(3.20) \quad \Omega_t(i, \xi) = \Omega_t(i) \cap \Omega_{t,\xi}.$$

It follows that $\Omega_t = \bigcup_{\xi \in R} \Omega_{t,\xi} = \bigcup_{\xi \in R} \bigcup_{i=t}^k \Omega_t(i, \xi)$. By using location invariance arguments similar to the scale invariance arguments of Lemma 3.1 it can be shown that

$$\text{Lemma 3.2.} \quad \inf_{\Omega_t} P[CS|R, (s, n)] = \min_{i \in I} \inf_{\Omega_t(i, 0)} P_{F_t}[A_i]$$

where A_i and I are defined as in subsection 3.2.

From the stochastic bounds developed in Corollary 2.2 for $F \in F_t(G)$ and arguments similar to those preceding Theorem 3.1 a tight lower bound, F_n^t , can be constructed for $\inf_{\Omega_t} P[CS|R_1(s, n)]$.

$$\text{Theorem 3.2.} \quad \inf_{\Omega_t} P[CS|R_1(s, n)] = P_{F(t)}[A_t] = F_n^t = F_n^t(r, t, s, k, \alpha, d^*)$$

where $F(t)$ is any configuration with t components $F_{t,s}(\cdot|s)$ and $(k-t)$ components $F_{t,\ell}(\cdot|-d^*)$ and

$$\begin{aligned}
 (3.21) \quad F_n^t &\equiv \int_{G(-d^*)}^{\alpha} H(G(G^{-1}(y) + d^*)) dJ(y) \\
 &\quad + H(G(d^*)) [1 - J(\alpha)].
 \end{aligned}$$

Remark 3.4 The proof of Theorem 3.2 consists in constructing a sequence of configurations in Ω_t such that the corresponding sequence of probabilities of correct selection converge to F_n^t . It can also be shown that $F_n^t \rightarrow 1$ as $n \rightarrow \infty$ and hence (3.8) can be guaranteed for any $p^* < 1$.

4. Random Size Subset Selection Procedures

4.1 Preferred Population Formulation

When the goal of an experiment is to screen a large number of populations to obtain a more manageable subset containing at least r preferred populations then the procedure $R_1(s, n)$ has the undesirable characteristic of always selecting a subset of size s . This section studies the random size restricted subset selection procedures introduced by Gupta and Santner (1973) and Santner (1975). These rules are characterized by the properties: (1) the experimenter specifies an upper bound, say s , on the size of the selected subset and (2) the procedure is able to capitalize on configurations favorable to the experimenter by selecting fewer than the maximum number of $s \pi_i$'s.

Following Santner (1975), restricted subset selection procedures will be defined in terms of a sequence of functions $\Psi \equiv \{\psi_n\}_{n=1}^{\infty}$, $\psi_n: \Xi \rightarrow \Xi$ satisfying:

$$(4.1) \quad \forall \xi \in \Xi \text{ and } \forall n, \psi_n(\xi) > \xi,$$

$$(4.2) \quad \forall n, \psi_n \text{ is continuous and strictly increasing and}$$

$$(4.3) \quad \forall \xi \in \Xi, \psi_n(\xi) \rightarrow \xi \text{ as } n \rightarrow \infty.$$

Example 4.1. For $\Xi = (0, \infty)$ take $\psi_n(\xi) = \xi \rho^{e(n)}$ where $\rho \in (1, \infty)$ and $\{e(n)\}$ is any sequence of positive numbers converging to zero as $n \rightarrow \infty$.

Example 4.2. For $\Xi = \mathbb{R}$ take $\psi_n(\xi) = \xi + e(n)$ where $e(n)$ is as in the previous example.

For specified $\alpha \in (0, 1)$, and each s in $\{1, \dots, k\}$, Ψ satisfying (4.1)--(4.3) and $n \geq N_\alpha$ define the procedure

$R_2(s, \Psi, n)$: Take independent random samples of size n from each population and select $\pi_i \Leftrightarrow T_i \geq \max\{T_{[k-s+1]}, \psi_n^{-1}(T_{[k]})\}$ where $T_{[1]} \leq \dots \leq T_{[k]}$ are the ordered values of the sample α -quantiles T_1, \dots, T_k .

Design Requirement: Given $P^* \in (t/k, 1)$, $1 \leq t < k$, $1 \leq s \leq k-t+1$, Ψ and $0 < c^* < 1$ ($0 < d^* < \infty$) determine the smallest $n \geq N_\alpha$ so that

$$(4.4) \quad P_{\tilde{F}}[CS | R_2(s, \Psi, n)] \geq P^* \quad \forall \tilde{F} \in \Omega_* (\Omega_t)$$

where the event $[CS | R_2(s, \Psi, n)]$ occurs iff at least one preferred π_i is selected.

Remark 4.1. In this formulation π_i is preferred $\Leftrightarrow x_\alpha(F_i) >$

$c^* x_\alpha(F_{[k-t+1]}) (x_\alpha(F_{[k-t+1]}) - d^*)$ in the star (tail) ordered model.

However only the goal corresponding to $r = 1$ of Section 3 is discussed.

The monotonicity arguments used below in establishing the infimum of the PCS are invalid when $r > 1$. When $R_2(k-t+1, \Psi, n)$ selects $k-t+1$ populations a correct selection automatically occurs and when $s = 1$ the rule reduces to $R_1(1, n)$. The results in Sections 4.2 and 4.3 are valid for $1 \leq t < k$ and $1 < s < k-t+1$. The case $s = k-t+1$ can be easily analysed by straightforward arguments similar to those developed below.

4.2 Selection from Star Ordered Families

Given $\alpha \in (0, 1)$ choose G as in Section 3.2 and for $1 \leq t < k$ define Ω_* , $\Omega_*(i)$, $\Omega_{*, \xi}$ and $\Omega_*(i, \xi)$ by (3.6), (3.9), (3.10) and (3.11) respectively. Clearly $P_{\tilde{F}}[CS | R_2(s, \Psi, n)] = 1 \quad \forall \tilde{F} \in \Omega_k$. Let

$I \equiv \{t, \dots, k-1\}$ and for $i \in I$ and $F_{\sim} \in \Omega_{\sim}$ define $W_i = \max\{T_{(k-i+1)}, \dots, T_{(k)}\}$ and

$$(4.5) \quad B_i = \begin{cases} \{W_i \geq \text{at least } (k-s+1-i) \text{ of } T_{(1)}, \dots, T_{(k-i)}\}; \\ W_i \rho^{e(n)} \geq \max\{T_{(j)} \mid 1 \leq j \leq k-i\} & , t \leq i \leq k-s \\ W_i \rho^{e(n)} \geq \max\{T_{(j)} \mid 1 \leq j \leq k-i\} & , k-s < i \leq k-1 \end{cases}$$

where as before $T_{(i)}$ is the sample α -quantile with cdf $F_{[i]}$. Then $\forall i \in I$ and $F_{\sim} \in \Omega_{\sim}$, $P_F[CS | R_2(s, \Psi, n)] = P_{F_{\sim}}[B_i]$. Furthermore the events $\{B_i\}$ form a nondecreasing sequence and $P_{F_{\sim}}[B_i] = E_{F_{\sim}}[\chi_{B_i}(T)]$ is nonincreasing in any of $F_{[1]}, \dots, F_{[k-i]}$ and nondecreasing in any of $F_{[k-i+1]}, \dots, F_{[k]}$ under \prec_{st} .

Lemma 4.1. $\inf_{\Omega_{\sim}} P_F[CS | R_2(s, \Psi, n)] = \min_{i \in I} \inf_{\Omega_{\sim}(i, 1)} P[B_i]$.

Proof. It suffices to show $\inf_{\Omega_{\sim}(i)} P_{F_{\sim}}[B_i] = \inf_{\Omega_{\sim}(i, 1)} P_{F_{\sim}}[B_i]$. When $k-s+1 \leq i \leq k-1$ and $F_{\sim} \in \Omega_{\sim}(i, \xi)$ then

$$\begin{aligned} P_{F_{\sim}}[B_i] &= P_{F_{\sim}}[\max\{T_{(j)} \rho^{e(n)} / \xi \mid k-i+1 \leq j \leq k\} \geq \max\{T_{(j)} / \xi \mid 1 \leq j \leq k-i\}] \\ &= P_{F'}[B_i] \text{ where } F'(x) = (F(x\xi), \dots, F_k(x\xi)) \text{ is in } \Omega_{\sim}(i, 1) \end{aligned}$$

as noted in the proof of Lemma 3.1. The proof for the case $t \leq i \leq k-s$ is similar.

Now the events $\{A_i\}$ and $\{B_i\}$ possess similar monotonicity properties and so the argument previously used in the construction of the lower bound F_n^* remains valid and yields

$$(4.6) \quad \inf_{\Omega_*} P_F[CS|R_2(s, \Psi, n)] \geq P_{F(t)}[B_t] \equiv R_n^*(t, s, k, \alpha, c^*) = R_n^*, \text{ say,}$$

where $F(t)$ is defined in Section 3.2.

The lower bound will now be evaluated. Let $T_{[q]}$ be the q^{th} ordered sample α -quantile, $\mathcal{D} \equiv \{k-t+1, \dots, k\}$, $\bar{\mathcal{D}} \equiv \{1, \dots, k-t\}$, for $q > t$ let $\{S_v^q | 1 \leq v \leq \binom{k-t}{q-t}\}$ be the collection of all subsets of size $(q-t)$ from $\bar{\mathcal{D}}$ and let $\bar{S}_v^q = \bar{\mathcal{D}} - S_v^q$. Then

$$\begin{aligned} R_n^* &= \sum_{q=k-s+1}^k P_{F(t)}[W_t = T_{[q]}; B_t] \\ &= \sum_{q=k-s+1}^k \sum_{j=k-t+1}^k P_{F(t)}[W_t = T_{(j)} = T_{[q]}; T_{(j)} \rho^{e(n)} \geq \max\{T_{(m)} | 1 \leq m \leq k-t\}] \\ &= \sum_{q=k-s+1}^k \sum_{j=k-t+1}^k \sum_{v=1}^{\binom{k-t}{q-t}} P_{F(t)} \left[\begin{array}{l} T_{(\ell)} < T_{(j)} \forall \ell \in S_v^q \cup \mathcal{D} \setminus \{j\} \\ T_{(j)} < T_{(\ell)} < T_{(j)} \rho^{e(n)} \forall \ell \in \bar{S}_v^q \end{array} \right] \\ &= \sum_{q=k-s+1}^k \binom{k-t}{q-t} \int_0^\infty \{B(F_{*,\ell}(y|c^*))\}^{q-t} \{D_n(F_{*,\ell}(y|c^*))\}^{k-q} d\{B(F_{*,s}(y|1))\}^t \\ &= \sum_{q=k-s+1}^k \binom{k-t}{q-t} \left[\int_{G(c^*)}^\alpha \{B(G(G^{-1}(y)/c^*))\}^{q-t} \{B(G(\frac{G^{-1}(y)\rho^{e(n)}}{c^*}))\} - \right. \\ &\quad \left. B(G(\frac{G^{-1}(y)}{c^*}))\}^{k-q} d\{B(y)\}^t \right. \\ &\quad \left. + \{B(G(\frac{1}{c^*}))\}^{q-t} \{D_n(G(\frac{1}{c^*}))\}^{k-q} [1 - B^t(\alpha)] \right] \end{aligned}$$

where $D_n(F(y)) = B(F(y)\rho^{e(n)}) - B(F(y))$. As in the case of $R_1(s, n)$

this lower bound is tight.

Theorem 4.1. $\inf_{\Omega_*} P_{F_{\sim}}[CS|R_2(s, \Psi, n)] = R_n^*$

The computation of $\lim_{j \rightarrow \infty} P_{F_{\sim j}}[CS|R_2(s, \Psi, n)]$ for the sequence $\{F_{\sim j}\}$ displayed in the proof of Theorem 3.1 yields the result. Furthermore $R_n^* \rightarrow 1$ as $n \rightarrow \infty$ and (4.4) can be guaranteed by finite n for any $P^* < 1$.

Example 4.3. When Ω_* is the class of all continuous IFRA distributions $(G(x) = 1 - e^{x \ln(1-\alpha)})$ then $R_n^*(1, s, k, \alpha, c^*)$ becomes

$$(4.7) \quad \sum_{q=k-s+1}^k \binom{k-1}{q-1} \left[\int_{1-(1-\alpha)}^{\alpha} c^{* \{B(1-(1-x)^{1/c^*})\}^{q-1} \{B(1-(1-x)^{\rho^{e(n)}/c^*}) - B(1-(1-x)^{1/c^*})\}^{k-q}} \right. \\ \left. dB(x) + \{B(1-(1-\alpha)^{1/c^*})\}^{q-1} \{1 - B(\alpha)\} \times \right. \\ \left. \{B(1-(1-\alpha)^{\rho^{e(n)}/c^*}) - B(1-(1-\alpha)^{1/c^*})\}^{k-q} \right]$$

Tables of the smallest odd sample sizes required to satisfy (4.4) were computed from (4.7) based on $\rho = 4$ and $e(n) = 1/\sqrt{n}$ for $\alpha = .25, .50, .75$, $c^* = .65, .70$, $P^* = .75, .90, .95$, $k = 3(1)9$ and $s = 2(1)\min\{4, k-1\}$.

4.3 Selection from Tail Ordered Families

This section studies the design problem for $R_2(s, \Psi, n)$ when $F_{\sim} \in \Omega_t$ and $\Psi_n(\xi) = \xi - e(n)$ where $\{e(n)\}$ is a sequence of positive numbers decreasing to zero. The developments follow from modifications of the arguments of Section 4.2 along the lines of Section 3.3 and hence only the final results will be stated.

		$c^* = .65$			$c^* = .70$		
		p^*			p^*		
k	s	.75	.90	.95	.75	.90	.95
3	2	27	71	107	43	103	155
4	3	31	75	115	47	111	167
4	2	35	83	127	55	123	183
5	4	35	83	127	51	123	183
5	3	39	87	127	55	123	187
5	2	43	95	139	65	139	203
6	4	39	91	135	59	131	195
6	3	43	95	139	63	135	199
6	2	51	107	151	75	155	223
7	4	43	99	143	67	143	207
7	3	47	103	147	69	147	211
7	2	57	115	163	85	167	235
8	4	47	103	151	71	151	219
8	3	51	107	155	75	159	223
8	2	63	123	171	93	179	251
9	4	51	107	155	77	159	227
9	3	55	115	163	83	167	235
9	2	67	131	179	99	191	263

Table 4.1. Sample sizes required by $R_2(s, \Psi, n)$ when selecting from Ω_* when $G(x) = 1 - e^{-x}$ with $\rho = 4$, $e(n) = 1/\sqrt{n}$ and $\alpha = .25$.

k	s	$c^* = .65$			$c^* = .70$		
		p^*			p^*		
		.75	.90	.95	.75	.90	.95
3	2	11	31	49	17	45	71
4	3	11	31	49	17	47	73
4	2	15	39	59	23	55	83
5	4	13	33	53	19	49	77
5	3	13	35	55	21	53	79
5	2	19	45	65	29	65	95
6	4	15	37	57	21	55	83
6	3	17	39	61	25	59	87
6	2	23	51	73	35	73	105
7	4	17	39	61	25	59	89
7	3	19	43	65	29	63	95
7	2	27	55	79	39	81	113
8	4	19	43	63	27	63	93
8	3	21	47	69	31	69	101
8	2	31	59	83	45	87	121
9	4	19	45	67	29	67	97
9	3	23	51	73	35	75	107
9	2	33	63	87	49	91	127

Table 4.2. Sample sizes required by $R_2(s, \Psi, n)$ when selecting from Ω_* when $G(x) = 1 - e^{-x}$ with $\rho = 4$, $e(n) = 1/\sqrt{n}$ and $\alpha = .50$.

k	s	c* = .65			c* = .70		
		p*			p*		
		.75	.90	.95	.75	.90	.95
3	2	7	23	35	11	31	51
4	3	7	23	35	11	31	51
4	2	11	27	43	17	39	59
5	4	7	23	35	11	35	51
5	3	11	27	39	15	35	55
5	2	15	33	47	21	47	69
6	4	11	27	39	15	35	55
6	3	11	27	43	17	41	63
6	2	17	37	53	25	53	77
7	4	11	27	43	15	39	59
7	3	13	31	47	19	45	67
7	2	21	41	57	29	59	83
8	4	13	31	43	19	43	63
8	3	15	33	49	21	49	71
8	2	21	43	61	33	63	89
9	4	13	31	47	21	47	67
9	3	17	37	53	25	53	75
9	2	25	47	65	37	67	93

Table 4.3. Sample sizes required by $R_2(s, \Psi, n)$ from Ω_{**} when $G(x) = 1 - e^{-x}$ with $\rho \approx 4$, $e(n) = 1/\sqrt{n}$ and $\alpha = .75$.

Theorem 4.2. $\inf_{\Omega_t} P[CS|R_2(s, \Psi, n)] = P_{F(t)}[B_t] = R_n^t$ where $F(t)$ is as in Theorem 3.2 and $R_n^t = R_n^t(t, s, k, \alpha, d^*)$ is given by

$$\sum_{q=k-s+1}^k \binom{k-t}{q-t} \left[\int_{G(-d^*)}^{\alpha} \{B(G[G^{-1}[x]+d^*])\}^{q-t} \{B(G[G^{-1}[x]+e(n)+d^*]) - B(G[G^{-1}[x]+d^*])\}^{k-q} d\{B(x)\}^t \right. \\ \left. + \{B(G[d^*])\}^{q-t} \{B(G[e(n)+d^*]) - B(G[d^*])\}^{k-q} \{1-B^t(\alpha)\} \right]$$

Furthermore $R_n^t \rightarrow 1$ as $n \rightarrow \infty$ and so (4.4) can be guaranteed for any $p^* < 1$.

Example 4.4. If Ω_t consists of all configurations with continuous components having "lighter" tails than the logistic distribution then $G(x) = [1 + \exp(-(x - \ln[(1-\alpha)/\alpha]))]^{-1}$. The lower bound R_n^t on the probability of correctly selecting at least one π_i having α -quantile within d^* of the largest α -quantile ($t=1$) is

$$\sum_{q=k-s+1}^k \binom{k-1}{q-1} \left[\int_{\frac{\alpha}{\alpha+(1-\alpha)e^{-d^*}}}^{\alpha} \{B(\frac{x}{x+(1-x)e^{-d^*}})\}^{q-1} \{B(\frac{x}{x+(1-x)e^{-[d^*+e(n)]}}) - B(\frac{\alpha}{\alpha+(1-\alpha)e^{-d^*}})\}^{k-q} dB(x) \right. \\ \left. + \{B(\frac{\alpha}{\alpha+(1-\alpha)e^{-d^*}})\}^{q-1} \{1-B(\alpha)\} \times \{B(\frac{\alpha}{\alpha+(1-\alpha)e^{-[d^*+e(n)]}}) - B(\frac{\alpha}{\alpha+(1-\alpha)e^{-d^*}})\}^{k-q} \right]$$

5. Properties of $R_1(s,n)$ and $R_2(s,\Psi,n)$

The purpose of this section is to describe some small sample performance characteristics of $R_1(s,n)$ and $R_2(s,\Psi,n)$ including (1) the effect of skewness and kurtosis of $R_1(s,n)$, (2) their monotonicity properties, (3) their performance as $c^* \rightarrow 1$ and some large sample properties including (1) a study of the number of populations selected by $R_2(s,\Psi,n)$ and (2) a proposal for choosing the sequence Ψ .

The first set of results show that the more skewed or the heavier the tails of the component F_i 's the smaller the probability of correct selection using $R_1(s,n)$. The analysis follows Doksum (1969) who used weak orderings to study the effects of skewness and kurtosis on the power of monotone rank tests. It has previously been established in Remark 3.2 that the "closer together" the preferred and nonpreferred populations are stochastically, the smaller the PCS. In studying the effects of skewness and kurtosis it is necessary to eliminate this stochastic effect by restricting attention to configurations of distributions differing only in scale or location. In this set up the resulting scale or location parameters are measures of "stochastic distance." For $1 \leq t < k$ and

$i \in I = \{i | t \leq i < r+k-s\}$ let $\mathcal{D}_i \equiv \{\Delta_i = (\delta_1, \dots, \delta_k) | \infty > \delta_1 \geq \delta_2 \geq \dots \geq \delta_{k-i} = 1 > \delta_{k-i+1} \geq \dots \geq \delta_k > 0 \text{ and } c^* \geq \delta_{k-t+1} > c^* \delta_{k-i+1}\}$ and for $F \in F_0$ and $\Delta_i \in \mathcal{D}_i$ let $F(\Delta_i) = (F_{\delta_1}, F_{\delta_2}, \dots, F_{\delta_k})$ where $F_{\delta_j}(x) = F(\delta_j x)$.

Remark 5.1. If $F \in F_0$ has $F(\xi) = \alpha$ then $\chi_\alpha(F_{\delta_j}) = \xi/\delta_j$. Hence $F(\Delta_i)$ contains exactly i preferred populations and $F_{[j]} = F_{\delta_j}$.

Theorem 5.1 If $F, H \in F_0$ satisfy $F \prec_{**} H$ and either $H(x) < 1 \forall x \in R$ or $F(x) < 1 \forall x \in R$ then $\forall i \in I$ and $\Delta_i \in \mathcal{D}_i$

$$P_{F(\Delta_i)}^{CS|R_1(s,n)} \geq P_{H(\Delta_i)}^{CS|R_1(s,n)}$$

Proof. Suppose $T_{(j)}$ has cdf $B(F_{\delta_j})$, $1 \leq j \leq k$, and define

$$(5.1) \quad T'_{(j)} = [B(H)]^{-1} B(F(T_{(j)})) = H^{-1}F(T_{(j)}) \quad \text{and}$$

$$(5.2) \quad T''_{(j)} = [B(H_{\delta_j})]^{-1} B(F_{\delta_j}(T_{(j)})) = H^{-1}F(\delta_j T_{(j)})/\delta_j.$$

It follows that $T'_{(j)}$ has cdf $B(H_{\delta_j})$ for $1 \leq j \leq k$ and $T'_{(k-i)} = T''_{(k-i)}$ has cdf $B(H)$ since $\delta_{k-i} = 1$.

Since $\delta_j \in (0,1)$ for $k-i+1 \leq j \leq k$ it follows from Lemma 2.2 that

$$(5.3) \quad T'_{(j)} = H^{-1}F(T_{(j)}) \geq H^{-1}F(\delta_j T_{(j)})/\delta_j = T''_{(j)}$$

provided $T_{(j)}$ and $\delta_j T_{(j)} \in R^+(F)$. Now $\delta_j T_{(j)}$ has distn $B(F)$ hence $\delta_j T_{(j)} \in R^+(F)$ a.s. (F_{δ_j}) . If $F(x) < 1 \forall x < \infty$ then $T_{(j)} \in R^+(F) = (0, \infty)$ while if $H(x) < 1 \forall x < \infty$ then $T_{(j)} \notin R^+(F) \Rightarrow F(T_{(j)}) = 1 \Rightarrow T'_{(j)} = H^{-1}(1) = \infty$. In either case (5.3) holds a.s. (F_{δ_j}) . Since $1 \leq \delta_j$ for $1 \leq j \leq k-i$ it also follows from $0 \leq T_{(j)} \leq \delta_j T_{(j)}$ that

$$(5.4) \quad T'_{(j)} = H^{-1}F(T_{(j)}) \leq H^{-1}F(\delta_j T_{(j)})/\delta_j = T''_{(j)}$$

holds a.s. (F_{δ_j}) .

Because $\chi_{A_i}(\cdot)$ depends only on the ranks of the $T_{(i)}$'s and $H^{-1}F(\cdot)$ is strictly increasing on $I(F)$ and nondecreasing on $(0, \infty)$ it can be seen that

$$(5.5) \quad X_{A_i}(T_{(1)}, \dots, T_{(k)}) = X_{A_i}(T'_{(1)}, \dots, T'_{(k)}) \text{ hence}$$

$$(5.6) \quad X_{A_i}(T_{(1)}, \dots, T_{(k)}) \geq X_{A_i}(T'_{(1)}, \dots, T'_{(k)})$$

holds a.s. $(F_{\tilde{A}_i})$ from (5.3), (5.4) and (5.5). Taking expectations of both sides of (5.6) wrt $F_{\tilde{A}_i}$ and recalling that $T'_{(j)}$ has cdf $B(H_{\delta_j})$ completes the proof.

Remark 5.2. If $R^+(F) = (0, b_F)$ and $R^+(H) = (0, b_H)$ where $\max(b_F, b_H) < \infty$ then (5.3) need not hold for $\delta_j \in (0, 1)$ and hence the proof is not valid in this case.

The results describing the effects of kurtosis on $P_F[CS|R_1(s, n)]$ will now be stated. For $F, H \in \mathcal{F}$ recall that $F \prec_t H$ means that H has heavier tails than F . For $\theta \in \mathbb{R}$ and $F \in \mathcal{F}$ let $F_\theta(x) = F(x + \theta)$ and for $i \in I$ define $F_i = \{\theta_i = (\theta_1, \dots, \theta_k) | \infty > \theta_1 \geq \theta_2 \geq \dots \geq \theta_{k-i} = 0 > \theta_{k-i+1} \geq \dots \geq \theta_k > -\infty; -d^* \geq \theta_{k-t+1} > \theta_{k-i+1} - d^*\}$ and for $F \in \mathcal{F}$ and $\theta_i \in F_i$ let $F_{\tilde{\theta}_i} = (F_{\theta_1}, \dots, F_{\theta_k})$. The configuration $F_{\tilde{\theta}_i}$ contains exactly i α^* -preferred populations. The following theorem shows the heavier the tails of the population distributions the smaller the probability of a correct selection.

Theorem 5.2. If $F, H \in \mathcal{F}$ satisfy $F \prec_t H$ and either $F(x) < 1 \forall x \in \mathbb{R}$ or $H(x) < 1 \forall x \in \mathbb{R}$ then $\forall i \in I$ and $\theta_i \in F_i$

$$P_{F_{\tilde{\theta}_i}}[CS|R_1(s, n)] \geq P_{H_{\tilde{\theta}_i}}[CS|R_1(s, n)].$$

Remark 5.3. The proofs of Theorems 5.1 and 5.2 are not valid for the procedure $R_2(s, \Psi, n)$ because the characteristic functions of the events B_i do not depend only on the ranks of the $T_{(j)}$'s.

The monotonicity properties of $R_1(s,n)$ and $R_2(s,\Psi,n)$ will next be investigated. Fix any $r, 1 \leq r \leq \min\{t,s\}$ when referring to $R_1(s,n)$ and $r=1$ when referring to $R_2(s,\Psi,n)$. Given an arbitrary selection procedure P and any integer i satisfying $r \leq i \leq k$ let

$$(5.7) \quad P_{\tilde{F}}[S_i|P] = P_{\tilde{F}}[P \text{ selects at least } r \text{ of } \pi(j_1), \dots, \pi(j_i)]$$

for any $S_i \equiv \{j_1, \dots, j_i\} \subset \{1, \dots, k\}$ of size i . For a given S_i let $j_{[1]} < \dots < j_{[i]}$ denote the ordered components of S_i .

Definition 5.1. P is monotone wrt Ω iff $\forall i, r \leq i \leq k$ and $\tilde{F} \in \Omega$

$$P_{\tilde{F}}[S_i|P] \leq P_{\tilde{F}}[S'_i|P] \text{ whenever } j_{[\ell]} \leq j'_{[\ell]} \quad \forall \ell, 1 \leq \ell \leq i.$$

Recall that for a given set of configurations $\Omega, \rho(\cdot)$ satisfying (3.1)-(3.4) and $t < k$ that $\Omega(i), t \leq i \leq k$, is the subset of Ω containing exactly i preferred populations.

Definition 5.2. P is weakly monotone wrt Ω iff $\forall i, t \leq i \leq k$ and

$$\tilde{F} \in \Omega(i), P_{\tilde{F}}[S_i|P] \leq P_{\tilde{F}}[S'_i|P] \text{ whenever } S_i - (S_i \cap S'_i) \subset \{1, \dots, k-i\} \text{ and } S'_i - (S_i \cap S'_i) \subset \{k-i+1, \dots, k\} \text{ where } - \text{ denotes set subtraction.}$$

Remark 5.4. The standard definitions of monotonicity used in parametric problems differ from those given here. In particular Definition 4.1 of monotonicity in Santner (1975) is a special case ($i = r = 1$) of the present definition. Also note that monotonicity implies weak monotonicity.

Lemma 5.1 Any procedure P in $\{R_1(s,n) | n \geq N_\alpha\}$ or in $\{R_2(s,\Psi,n) | n \geq N_\alpha\}$ must satisfy $P_{\tilde{F}}[S_i|P] \leq P_{\tilde{F}'}[S_i|P] \quad \forall i, r \leq i \leq k$ and $\tilde{F}, \tilde{F}' \in \Omega(F) = F^k \Rightarrow F[j] \prec_{st} F'[j] \quad \forall j \in S_i$ and $F'[j] \prec_{st} F[j] \quad \forall j \in \bar{S}_i \equiv \{1, \dots, k\} - S_i$.

The result is immediate since $[S_i | R_1(s, n)] = [R_1(s, n)]$ selects at least r of $\pi_{(j)}, j \in S_i$ and $[S_i | R_2(s, \Psi, n)] = [R_2(s, \Psi, n)]$ selects at least one of $\pi_{(j)}, j \in S_i$ are nondecreasing in any of $T_{(j)}, j \in S_i$ and nonincreasing in any of $T_{(j)}, j \in \bar{S}_i$.

Some additional notation will now be introduced. Let

$$\Omega_{st} = \{F \in \Omega(F) | F_{[1]} \prec_{st} F_{[2]} \prec_{st} \dots \prec_{st} F_{[k]}\},$$

$$\Omega_R(i) = \left\{ \begin{array}{l} \{F \in \Omega(F) | x_\alpha(F_{[k-i]}) \leq p(x_\alpha(F_{[k-t+1]})) < x_\alpha(F_{[k-i+1]}); \\ \min\{F_{[\ell]}(y) | 1 \leq \ell \leq k-i\} \geq \max\{F_{[\ell]}(y) | k-i < \ell \leq k\} \forall y \in R\} \\ \text{for } t \leq i \leq k-1 \\ \{F \in \Omega(F) | p(x_\alpha(F_{[k-t+1]})) < x_\alpha(F_{[1]})\} \text{ for } i=k, \end{array} \right.$$

and $\Omega_R = \bigcup_{i=t}^k \Omega_R(i)$. Note that $\Omega_{st} \subset \Omega_R$.

Theorem 5.3. The classes of procedures $\{R_1(s, n) | n \geq N_\alpha\}$ and $\{R_2(s, \Psi, n) | n \geq N_\alpha\}$ are all monotone wrt any $\Omega_1 \subset \Omega_{st}$ and weakly monotone wrt any $\Omega_2 \subset \Omega_R$.

Proof. To show monotonicity it suffices to consider the following case.

Fix $F \in \Omega_1$, $r \leq i \leq k$, $S_i \subset \{1, \dots, k\}$, $j \in S_i$ and $m > j$ with $m \in \bar{S}_i$ and let $S'_i = S_i \cup \{m\} - \{j\}$. Lemma 5.1 implies

$$\begin{aligned}
(5.8) \quad P_{\tilde{F}}[S_i | R_1(s, n)] &= P[S_i | R_1(s, n)] \\
&\quad (F_{[1]}, \dots, F_{[j]}, \dots, F_{[m]}, \dots, F_{[k]}) \\
&\leq P[S_i | R_1(s, n)] \\
&\quad (F_{[1]}, \dots, F_{[m]}, \dots, F_{[m]}, \dots, F_{[k]}) \\
&= P[S_i' | R_1(s, n)] \\
&\quad (F_{[1]}, \dots, F_{[m]}, \dots, F_{[m]}, \dots, F_{[k]}) \\
&\leq P_{\tilde{F}}[S_i' | R_1(s, n)]
\end{aligned}$$

Since $R_2(s, \Psi, n)$ satisfies Lemma 5.1 the same proof also shows that it is monotone wrt Ω_1 . The proofs that $R_1(s, n)$ and $R_2(s, \Psi, n)$ are weakly monotone wrt Ω_2 are similar.

Remark 5.5. In most parametric selection problems where the scalar of interest, say λ_i , is estimated by T_i having cdf $F(\cdot | \lambda_i)$ the following holds: $\lambda_i < \lambda_j \Rightarrow F(\cdot | \lambda_i) <_{st} F(\cdot | \lambda_j)$ and hence the stronger property of monotonicity holds. However $x_\alpha(F_i) < x_\alpha(F_j) \not\Rightarrow F_i <_{st} F_j$ and hence the property of weak monotonicity is introduced here. An example will be given later to show that neither $R_1(s, n)$ nor $R_2(s, \Psi, n)$ is even weakly monotone wrt Ω_* or Ω_t .

Remark 5.6. Since $F_I \subset F$ for any interval $I \subset R$, Theorem 5.3 holds if Ω_* and Ω_k are defined as the appropriate subsets of F_I^k rather than F^k .

The next result, obtainable from straightforward computations, describes the behavior of the infimum of $P[CS | R_1(s, n)]$ over Ω_* as c^* increases to 1.

Lemma 5.2. $\lim_{c^* \rightarrow 1 - \Omega_*} \inf P[CS | R_1(s, n)] = B[B(\alpha); k-t-s+r, k-t] \times$

$$(1 - B[B(\alpha); t-r+1, t]) < (k-t-s+r-1) \binom{k-t}{s-j} \sum_{j=r}^t \binom{t}{j-r} / ((k-t-s+j-1) \binom{k}{s+t-j})$$

When $r=t$ the right hand side reduces to $\binom{k-t}{s-t}/\binom{k}{s}$ which is the probability of making a correct selection by randomly selecting s populations when exactly t are preferred. So for any fixed sample size $n \geq N_\alpha$ there exists c^* sufficiently close to 1 and an $F_\gamma \in \Omega_{**}(t)$ for which $P_{F_\gamma}[CS|R_1(s,n)] < P_{F_\gamma}[CS | \text{choose } s \text{ populations at random}]$. To gain an intuitive feel for this result choose an arbitrary $F_\gamma \in \Omega_{**}(t,1)$. Then $\forall x \in [c^*, 1]$

$$(5.9) \quad F_{[j]}(x) > F_{[i]}(x) \quad \forall 1 \leq j \leq k-t \quad \text{and} \quad k-t+1 \leq i \leq k.$$

However (5.9) need not hold for $x \notin [c^*, 1]$. In fact given $\epsilon > 0$ there exists $F_\gamma = F_\gamma(\epsilon) \in \Omega_{**}(t,1)$ such that for all $x \notin [c^* - \epsilon, 1 + \epsilon]$

$$(5.10) \quad F_{[j]}(x) < F_{[i]}(x) \quad \forall 1 \leq j \leq k-t \quad \text{and} \quad k-t+1 \leq i \leq k.$$

In such a case unless n is sufficiently large to detect the situation in $[c^*, 1]$ the procedure $R_1(s,n)$ will tend to reflect (5.10) and choose populations in $\{\pi_{(1)}, \dots, \pi_{(k-t)}\}$.

The preceding argument can be used to show that $R_1(t,n)$ is not weakly monotone wrt Ω_{**} . For fixed n and c^* sufficiently close to 1 choose $F_\gamma \in \Omega_{**}(t)$ such that

$$(5.11) \quad P_{F_\gamma}[CS|R_1(t,n)] = P_{F_\gamma}[k_t|R_1(t,n)] < 1/\binom{k}{t}$$

where $k_t = \{k-t+1, \dots, k\}$. Let $\{j(\ell) | 1 \leq \ell \leq \binom{k}{t}\}$ be the collection of all sets of size t from $\{1, \dots, k\}$. Then

$$\sum_{\ell=1}^{\binom{k}{t}} P_{\mathcal{F}}[j(\ell) | R_1(t,n)] = 1 \text{ and } (5.11) \Rightarrow \exists \text{ a } j^*(\ell) \text{ s.t.}$$

$$P_{\mathcal{F}}[j^*(\ell) | R_1(t,n)] > 1/\binom{k}{t} > P_{\mathcal{F}}[k_t | R_1(t,n)]$$

and hence $R_1(t,n)$ is not weakly monotone. Similar examples exist when Ω_* or $R_1(s,n)$ is replaced by Ω_t or $R_2(s,\Psi,n)$ respectively.

Some results will now be developed for the number of populations selected by $R_2(s,\Psi,n)$. Let $E(i,n) = [T_{(i)} \geq \max\{T_{[k-s+1]}, \psi_n^{-1}(T_{[k]})\}]$ be the event that $\pi(i)$ is selected, $P(i,n) = E_{\mathcal{F}}[E(i,n)]$ and $S(n) = \sum_{i=1}^k \chi_{E(i,n)}(T)$ the number of populations selected by $R_2(s,\Psi,n)$. It is straightforward to obtain an expression for $E_{\mathcal{F}}[S(n)]$ for arbitrary \mathcal{F} based on this representation for $S(n)$.

For all $F \in \mathcal{F}$ and $\alpha \in (0,1)$ it is well known that $T_{(i),n} \xrightarrow{D} \chi_{\alpha}(F)$ where $T_{(i),n}$ has cdf $B[F(y);q,n](n \geq N_{\alpha})$. The following results are a direct consequence of this fact and the proof of Theorem 5.2 and Corollaries 5.1 and 5.2 of Santner (1975).

Theorem 5.4. For any $F \in \Omega(F) \ni \chi_{\alpha}(F_{[k]}) > \chi_{\alpha}(F_{[k-1]})$

$$(1) \quad P_F(i,n) \rightarrow \begin{cases} 1 & , i = k \\ 0 & , 1 \leq i < k \end{cases} \text{ as } n \rightarrow \infty,$$

$$(2) \quad Z_i(n) \xrightarrow{L^2} \begin{cases} 1 & , i = k \\ 0 & , 1 \leq i < k \end{cases} \text{ as } n \rightarrow \infty \text{ and}$$

$$(3) \quad S(n) \xrightarrow{L^2} 1 \text{ and } E_{\mathcal{F}}[S(n)] \rightarrow 1 \text{ as } n \rightarrow \infty \text{ where } \xrightarrow{L^2} \text{ denotes convergence in the } L^2 \text{ norm.}$$

The final topic of this section is a proposal for choosing the constant $\rho \in (1,\infty)$ when the sequence $\Psi_n(\xi) = \xi \rho^{e(n)}$ is used to determine the rule $R_2(s,\Psi,n)$. The discussion is limited to the star ordered case but similar results hold for the tail ordered case.

Let $[NPS|R_2(s, \Psi, n)] = [R_2(s, \Psi, n)$ selects at least one non preferred population] and for $\rho \in (1, \infty)$ let $n(\rho)$ be the smallest $n \geq N_\alpha$ satisfying both

$$(5.12) \quad \inf_{\Omega_*} P_{\tilde{F}}[CS|R_2(s, \Psi, n)] \geq P^* \quad \text{and}$$

$$(5.13) \quad \sup_{\Omega_*} P_{\tilde{F}}[NPS|R_2(s, \Psi, n)] \leq \epsilon^*$$

for specified $\epsilon^* \in (0, 1)$ and $P^* \in (1/k, 1)$. The optimal choice of ρ minimizes $n(\rho)$. Arguments similar to those of Section 4.2 show the supremum of (5.13) occurs at $\tilde{F}(1, 1)$. Calculation gives

$$(5.14) \quad P_{\tilde{F}(1,1)}[NPS|R_2(s, \Psi, n)] = 1 - \int_0^\infty \{B(F_{*,\ell}(y/\rho^{e(n)}|c^*))\}^{k-1} dB(F_{*,s}(y|1))$$

$$= \begin{cases} 1 - \int_{c^* \rho^{1/\sqrt{n}}}^1 [B(G(y/c^* \rho^{1/\sqrt{n}}))]^{k-1} dB(G(y)) \\ \quad - [B(G(1/c^* \rho^{1/\sqrt{n}}))]^{k-1} [1 - B(\alpha)] & , c^* \rho^{1/\sqrt{n}} \geq 1 \\ 1 & , c^* \rho^{1/\sqrt{n}} < 1 \end{cases}$$

Remark 5.8. For fixed $\rho \in (1, \infty)$ the righthand side of (5.14) converges to zero as $n \rightarrow \infty$ and hence $\forall \epsilon^* \in (0, 1) \exists n \geq N_\alpha$ satisfying (5.13).

For each $\rho \in (1, \infty)$ let $n_1(\rho)$ be the smallest $n \geq N_\alpha$ satisfying (5.12) and let $n_2(\rho)$ be the smallest $n \geq N_\alpha$ satisfying (5.13). Since both $\sup_{\Omega_*} P_{\tilde{F}}[NPS|R_2(s, \Psi, n)]$ and $\inf_{\Omega_*} P_{\tilde{F}}[CS|R_2(s, \Psi, n)]$ are nondecreasing in ρ it follows that $n_1(\rho)$ is nonincreasing in ρ and $n_2(\rho)$ is nondecreasing in ρ . An optimal choice of ρ is ρ^* such that

$$(5.15) \quad \max\{n_1(\rho^*), n_2(\rho^*)\} = \min_{1 < \rho < \infty} \max\{n_1(\rho), n_2(\rho)\}.$$

The solution of (5.15) can be obtained via the same techniques employed in the construction of Tables 4.1 - 4.3 (see Hooper (1977)).

REFERENCES

- [1] Alam, K. and Rizvi, M.H. (1966). Selection from multivariate normal populations. Ann. Inst. Statist. Math. (18), 307-318.
- [2] Barlow, R.E. and Gupta, S.S. (1969). Selection procedures for restricted families of probability distributions. Ann. Math. Statist. (40), 905-917.
- [3] Barlow, R.E. and Proschan, F. (1975). Statistical Theory of Reliability and Life Testing, Holt, Rinehart and Winston, N.Y.
- [4] Bartlett, N.S. and Govindarajulu, Z. (1968). Some distribution-free statistics and their application to the selection problem. Ann. Inst. Statist. Math. (20), 79-97.
- [5] Bechhofer, R.E. (1954). A single-sample multiple decision procedure for ranking means of normal populations with known variances. Ann. Math. Statist. (25), 273-289.
- [6] Bhapkar, V.P. and Gore, A.P. (1971). Some selection procedures based on U-statistics for the location and scale problems. Ann. Inst. Statist. Math. (23), 375-86.
- [7] Carroll, R.J. (1974). Some Contributions to the Theory of Parametric and Nonparametric Sequential Ranking and Selection. Ph.D. Thesis, Dept. of Statist., Purdue University, West Lafayette, Indiana 47907.
- [8] Bruckner, A.M. and Ostrow, E. (1962). Some function classes related to the class of convex functions. Pacific J. Math. (12), 1203-1215.
- [9] Birnbaum, Z.W., Esary, J.D. and Marshall, A.W. (1966). A stochastic characterization of wear-out for components and systems. Ann. Math. Statist. (37), 816-825.
- [10] Desu, M. and Sobel, M. (1971). Nonparametric procedures for selecting fixed size subsets. In Statistical Decision Theory and Related Topics (ed. S.S. Gupta and J. Yackel), Academic Press, NY 255-274.
- [11] Doksum, K. (1969). Starshaped transformations and the power of rank tests. Ann. Math. Statist. (40), 1167-1176.
- [12] Eaton, Morris L. (1967). Some optimum properties of ranking procedures. Ann. Math. Statist. (36), 124-137.
- [13] Esary, J.D., Marshall, A.W. and Proschan, F. (1973), Shock models and wear processes. Ann. Probability (1), 627-649.
- [14] Ghosh, M. (1973). Nonparametric selection procedures for symmetric location parameter problems. Ann. Statist. (1), 773-779.

- [15] Gupta, S.S. (1956). On a decision rule for a problem in ranking means. Mimeograph Series No. 150, Institute of Statistics, Univ. of North Carolina, Chapel Hill.
- [16] Gupta, S.S. (1965). On some multiple decision (selection and ranking) rules. Technometrics (7), 225-245.
- [17] Gupta, S. S. and Huang, D.Y. (1974). Nonparametric subset selection procedures for the t Best Populations. Bulletin of the Institute of Mathematics Academia Sinica (2), 377-386.
- [18] Gupta, S.S. and McDonald, G.C. (1970). On some classes of selection procedures based on ranks. Nonparametric Techniques in Statistical Inference (ed. M.L. Puri), University of Cambridge Press, Cambridge.
- [19] Gupta, S.S. and Santner, T.J. (1973). On selection and ranking procedures--a restricted subset selection rule. Proceedings of the 39th Session of the International Statistical Institute, Vienna, Austria (1), 409-17.
- [20] Hooper, J.H. (1977). Selection procedures for ordered families of distributions. Ph.D. Thesis, School of Operations Research and Industrial Engineering, Cornell University.
- [21] Lee, Y.J. and Dudewicz, E.J. (1974). Nonparametric ranking and selection procedures. Technical Report No. 105, Department of Statistics, Ohio State Univ.
- [22] Lehmann, E.L. (1963). A class of selection procedures based on ranks. Math. Annalen (150), 268-275.
- [23] Mahamunala, D.M. (1967). Some fixed-sample ranking and selection problems. Ann. Math. Statist. (38), 1079-1091.
- [24] Panchapakesan, S. (1969). A class of selection procedures including procedures for restricted families of distributions. Mimeograph series No. 189, Department of Statistics, Purdue Univ.
- [25] Patel, J.K. (1976). Ranking and selection of IFR populations based on means. J. Amer. Statist. Assoc. (71), 143-146.
- [26] Puri, M.L. and Sen, P.K. (1972). On some selection procedures in two-way layouts. Z. Wahr. verw. Geb. (22), 242-50.
- [27] Randles, R.H. (1970). Some robust selection procedures. Ann. Math. Statist. (41), 1640-45.
- [28] Rizvi, M.H. and Sobel, M. (1967). Nonparametric procedures for selecting a subset containing the population with the largest α -quantile. Ann. Math. Statist. (38), 1788-1803.
- [29] Rizvi, M.H., Sobel, M. and Woodworth, G.G. (1968). Nonparametric ranking procedures for comparison with a control. Ann. Math. Statist. (39), 2075-2093.

- [30] Santner, T.J. (1975). A restricted subset selection approach to ranking and selection problems. Ann. Statist. (3), 334-349.
- [31] Shampine, L. and Allen, R. (1973). Numerical Computing: an Introduction, W.B. Saunders Company, Philadelphia.
- [32] Sobel, M. (1967). Nonparametric procedures for selecting the t populations with the largest α -quantiles. Ann. Math. Statist. (38), 1804-1816.
- [33] van Zwet, W.R. (1964). Convex transformations of random variables. Mathematical Centre Tract 7, Mathematisch Centrum, Amsterdam.

1. REPORT NUMBER #350	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) DESIGN OF EXPERIMENTS FOR SELECTION FROM ORDERED FAMILIES OF DISTRIBUTION		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Jeffrey H. Hooper and Thomas J. Santner		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0586 DAAG29-77-C-0003
9. PERFORMING ORGANIZATION NAME AND ADDRESS School of Operations Research and Industrial Engineering, College of Engineering Cornell University, Ithaca, NY 14853		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Sponsoring Military Activity: U.S. Army Research Office, P.O. Box 12211 Research Triangle Park, North Carolina 27709		12. REPORT DATE August 1977
		13. NUMBER OF PAGES 50
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Sponsoring Military Activity: Statistics and Probability Program Office of Naval Research Arlington, Virginia 22217		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release, distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Star ordering, Tail ordering, Preferred Population Formulation, Fixed subset selection procedure, Restricted subset selection procedure		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Given $G \in S$ a space of cumulative distribution functions and a weak (reflexive and transitive) order relation $<$ on S , the subclass of S given by $\{F \in S F < G\}$ is called an ordered family of distributions. Suppose π_1, \dots, π_k represent k populations with distributions only		

known to belong to some specified ordered family. The general problem is to design an experiment to select π_i 's having large (small) α -quantiles. A preferred population is defined to be any population with α -quantile "near" the t^{th} largest α -quantile and a correct selection occurs if the subset of populations selected contains at least a prespecified number, r , of preferred populations. The design problem is solved for both fixed and random subset size selection procedures under star and tail ordering. Tables of the sample sizes required to guarantee prespecified minimum probabilities of correct selection are given for the case of selecting from continuous IFRA populations. Comparisons are made with the procedure used in selecting from exponential populations. Properties of the proposed rules are discussed.